AD_____

Award Number: W81XWH-05-1-0441

TITLE: Development of a Computational Assay for the Estrogen Receptor

PRINCIPAL INVESTIGATOR: George C. Shields, Ph.D.
Karl N. Kirschner, Ph.D.

CONTRACTING ORGANIZATION: Hamilton College
Clinton NY 13323

REPORT DATE: July 2006

TYPE OF REPORT: Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
01/07/06	Final	1 Jul 05 – 30 Jun 06

4. TITLE AND SUBTITLE
Development of a Computational Assay for the Estrogen Receptor

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-05-1-0441

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
George C. Shields, Ph.D. ; Karl N. Kirschner, Ph.D.

E-Mail: gshields@hamilton.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Hamilton College
Clinton NY 13323

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
Estrogen Receptor alpha (αER), binds estradiol, initiating a cascade of events that leads to αER-positive breast cancer. The αER binds the FDA approved drug for breast cancer, tamoxifen, and also raloxifene, with high affinity. The crystal structures of the αER-estradiol, αER-tamoxifen, and ER-raloxifene complexes reveal that estradiol binding changes the conformation of helix 12 (H12) so that co-regulatory proteins can bind and activate transcription, whereas tamoxifen and raloxifene binding displaces H12 to a different position that prevents the binding of co-regulatory proteins.
We have successfully combined MD & QM methods in an innovative way to develop a computational assay that will allow for the development of novel lead compounds that should be active against breast cancer. Our test system consists of estradiol, tamoxifen, and several known raloxifene analogs. MD simulations were performed for 3-5 nanoseconds on a truncated αER protein bound to different ligands. These simulations generated input structures for a variety of methods, ranging from MM to QM, for determining energy the binding energy of ligands to the αER. We have determined the optimal method for correlating binding energies to experimental relative binding affinitites (RBAs). We will use this assay to design novel inhibitors with high-predicted activity.

15. SUBJECT TERMS
Estrogen Receptor, Computational Assay, Raloxifene, Tamoxifen, Estradiol, AMBER

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRMC
U	U	U	UU	138	19b. TELEPHONE NUMBER *(include area code)*

Table of Contents

Body

Abstract

Estrogen Receptor alpha (αER), binds estradiol, initiating a cascade of events that leads to αER-positive breast cancer. The αER binds the FDA approved drug for breast cancer, tamoxifen, and also raloxifene, with high affinity. The crystal structures of the αER-estradiol, αER-tamoxifen, and αER-raloxifene complexes reveal that estradiol binding changes the conformation of helix 12 (H12) so that co-regulatory proteins can bind and activate transcription, whereas tamoxifen and raloxifene binding displaces H12 to a different position that prevents the binding of co-regulatory proteins.

We have successfully combined MD & QM methods in an innovative way to develop a computational assay that will allow for the development of novel lead compounds that should be active against breast cancer. Our test system consists of estradiol, tamoxifen, and several known raloxifene analogs. MD simulation were performed for 3-5 nanoseconds on a truncated αER protein bound to different ligands. These simulations generated input structures for a variety of methods, ranging from MM to QM, for determining energy the binding energy of ligands to the αER. We have determined the optimal method for correlating binding energies to experimental relative binding affinitites (RBAs). We will use this assay to design novel inhibitors with high-predicted activity.

Introduction

The estrogen receptor (ER), a member of the steroid receptor nuclear family, resides in the nucleus of cells in its inactive form. In the presence of estrogen it becomes activated and the resulting ER-estradiol complex forms a stable dimer that subsequently initiates the transcription of target genes.[1] The ER has been studied extensively and found to have a significant role on cholesterol and lipid levels, the skeletal system, the central nervous system, as well as reproductive functions in both males and females.[2-4]

The site at which estrogen binds to the ER is called the ligand binding domain (LBD).[5] The LBD is a small hydrophobic pocket that can accommodate ~450 Å.[3] Estrogen's volume is only 245 Å^3,[4] which means that the relatively large LBD should be able to accommodate and bind other hydrophobic ligands of greater size, and in fact, it does. When another ligand binds to the ER and mimics estrogen's physiological affects, it is called an agonist, and when it acts as an inhibitor, it is called an antagonist.

While estrogen-mediated transcription of genes is essential for health, some aggressive human breast and endometrial cancers exhibit this same mechanism for growth.[6] In fact, measurable amounts of ER can be found in 75% of breast tumors.[7] In light of these findings, a great deal of work has been done to characterize ligands that bind antagonistically to the ER in tumor cells, while behaving as an agonist in other normal tissues in order to minimize the antagonist activity. The ER ligands that display this selective binding profile are called selective estrogen receptor modulators (SERMs). SERMs have been used to effectively treat breast cancer and other estrogen-regulated disease.

The most popular SERM used clinically is tamoxifen. The drug itself has only a mild affinity for the αER, however it undergoes modification once ingested, which increases its affinity for αER by 100-fold. Tamoxifen encounters members of the cytochrome p450 enzyme (CYPs) superfamily in the liver as well as the intestine, where it is ultimately metabolized into two products, 4-hydroxytamoxifen and endoxifen, as seen in Figure 1, both of which are 100 times better at binding with αER than tamoxifen.[8]

It has been shown to inhibit breast cancer and to have several other health benefits.[9,10] Because tamoxifen acts like estrogen in bone, it reduces bone resorption and consequently leads to higher bone density.[9] Many studies have shown that, like estrogen, tamoxifen and other SERMs have neuroprotective properties. The mechanisms for this protection are not understood, but further investigation may illicit possible benefits on neurodegenerative diseases like Alzheimer's and Parkinson's.[11] However, recent studies have linked tamoxifen to dramatically increased risks of endometrial cancer. In 2004, Curtis et al. provided population based evidence that the use of tamoxifen increases a patient's relative risk for uterine corpus cancer more than twofold, and for malignant mixed

mullerian tumors by fourfold.[12]

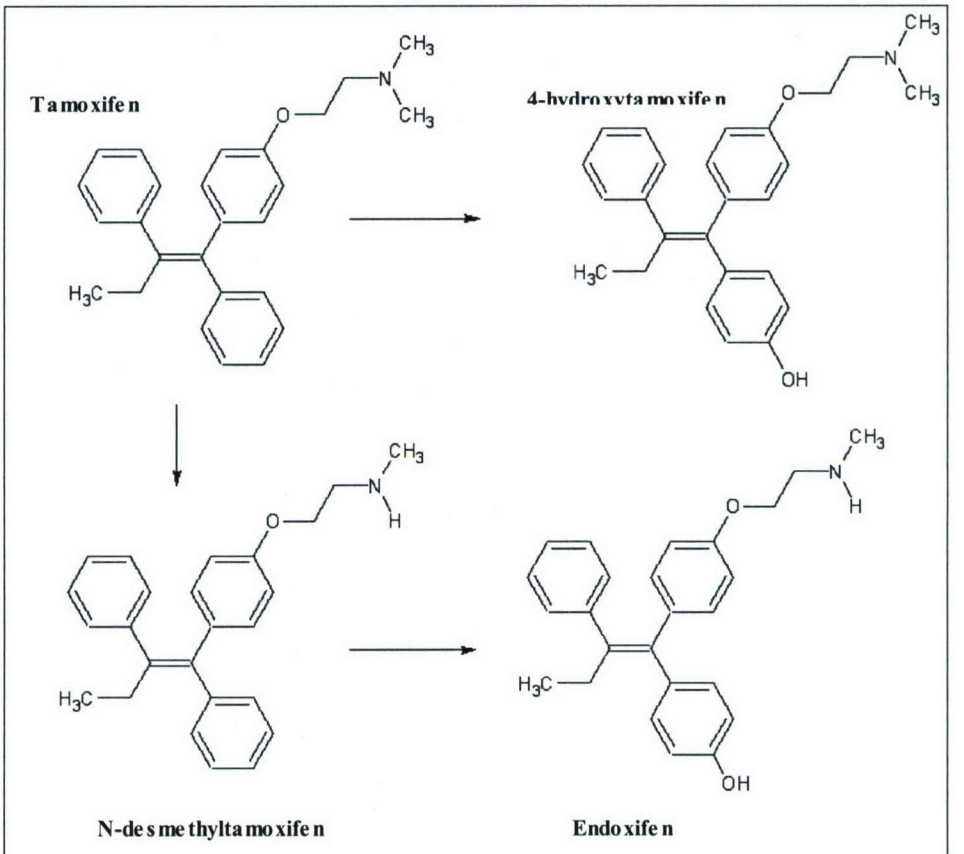


***Figure 1**: Metabolism of tamoxifen. Modifications made to tamoxifen while in the body make it better suited for its task of competitively inhibiting the ERα. Adapted from Holmes et al.[8]*

For this reason, researchers are looking for new drugs to treat breast cancer. A "second generation" SERM that has promising qualities is raloxifene, seen in Figure 2. It has a similar affinity for αER,[13] but acts as an antagonist both in breast and endometrial tissues.[14] In addition, it shares many of the same health benefits of tamoxifen.[15,16] It has been shown to inhibit breast cancer and to have several other health benefits.[9,10] During the Multiple Outcomes of Raloxifene Evaluation (MORE) trial, it was found that for women who received raloxifene, the risk of invasive breast cancer was reduced by 76% after 3 years and 72% after 4 years compared to the women who received placebo.[17,18] Because tamoxifen acts like estrogen in bone, it reduces bone resorption and consequently leads to higher bone density.[9] Many studies have shown that, like estrogen, tamoxifen and other SERMs have neuroprotective properties. The mechanisms for this protection are not understood, but further investigation may illicit possible benefits on neurodegenerative diseases like Alzheimer's and Parkinson's.[11] However, recent studies have linked tamoxifen to dramatically increased risks of endometrial cancer. In 2004, Curtis et al. provided population based evidence that the use of tamoxifen increases a patient's relative risk for uterine corpus cancer more than twofold, and for malignant mixed mullerian tumors by fourfold.[12]

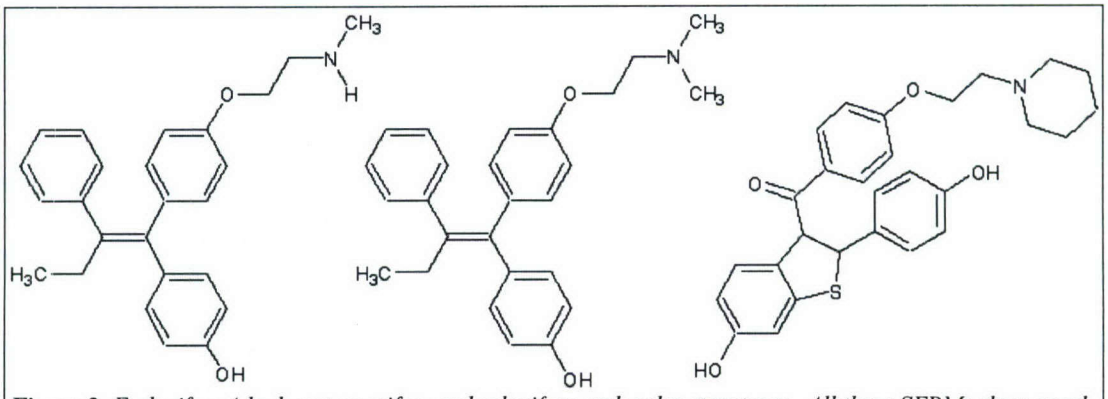


***Figure 2**: Endoxifen, 4-hydroxytamoxifen, and raloxifene molecular structures. All three SERMs show good binding affinity for αER, however the tamoxifen derivatives act as agonists in endometrial tissue, increasing risks of cancer while raloxifene acts as an antagonist in the uterus.*

Estrogen therapy in postmenopausal women has been associated with higher risk of coronary heart disease (CHD) as well as fatal strokes,[20] however raloxifene has not been associated with these risks. After the MORE trial, a *post hoc* analysis was done using data from the participants that determined that patients who received raloxifene were not at an increased risk of cardiovascular events when compared to those who received placebos. For these patients, who were considered high-risk for CHD, taking raloxifene actually reduced their risk of a cardiovascular incident. Currently, there is an international, randomized, double-blind, placebo-controlled study of over 10,000 postmenopausal women with increased risk of coronary heart disease called Raloxifene Use in The Heart (RUTH) that is underway which will determine whether raloxifene actually has significant cardiovascular benefit on patients.[21]

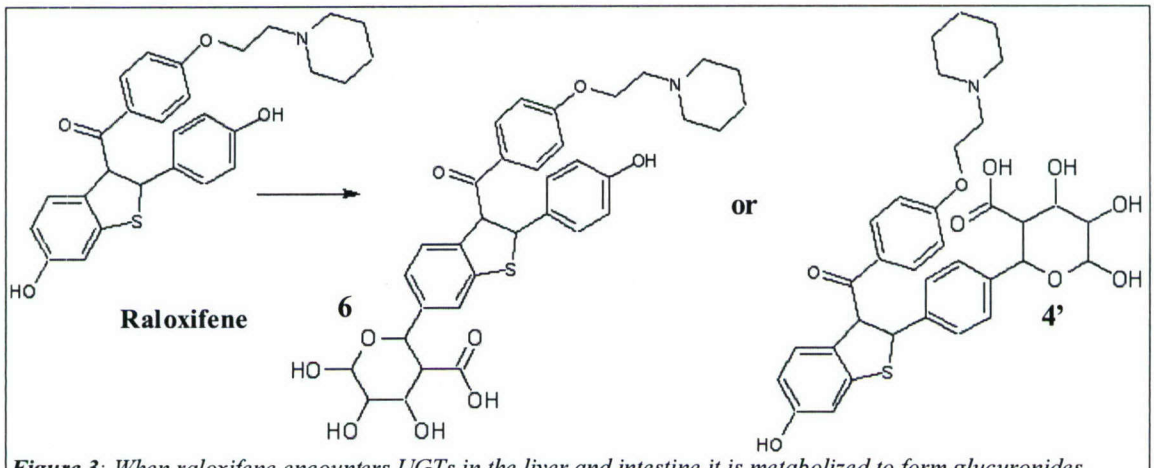


***Figure 3**: When raloxifene encounters UGTs in the liver and intestine it is metabolized to form glucuronides.*

While raloxifene's binding profile is highly desirable for cancer treatment, it is not an effective anti-cancer drug because only 2% of the dose becomes bioavailable.[22] Once it enters the body, raloxifene faces a host of obstacles. Sixty percent of the drug is rapidly absorbed after ingestion,[23] but then it undergoes rapid presystemic glucuronidation both in the liver and the intestine. The process of glucuronidation is one of the body's mechanisms for increasing substrates water solubility, but it acts as a significant barrier in the way of drug development. Research has shown that raloxifene is susceptible to this process primarily at the 4' position, but also at the 6 position, shown in Figure 3. The UDP-glucuronsyltransferases (UGTs) in the liver, primarily UGT1A1, catalyzes glucuronidation primarily at the 6' position while intestinal UGTs, UGT1A8 and UGT1A10, make 4' glucuronides preferentially.[22]

The rate of clearance of metabolism is surprisingly higher in humans than in rats, considering present in human intestine, which ultimately explains why, in rats, there is a 39% bioavailability compared to the human 2%. These findings also support the idea that intestinal enzymes play a more critical role in the metabolism of raloxifene than those in the liver.[24] While glucuronidation seems to be the major obstacle in the way of bioavailability, it has also been shown that raloxifene has difficulties with transport and is effluxed by several transport proteins, such as multidrug resistance related protein (MRP) and organic anion transporter (OAT).[25]

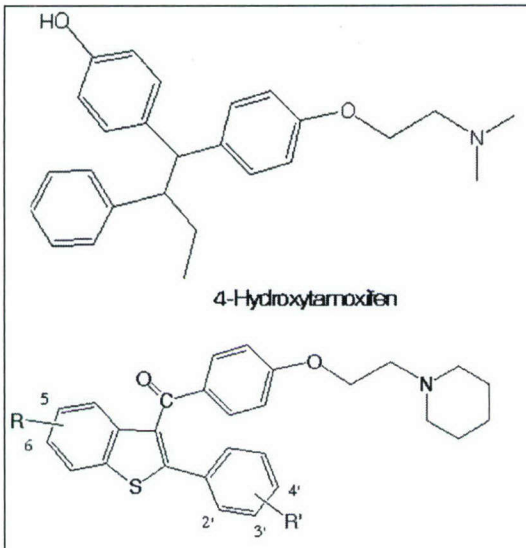
One of the aspects of ER-ligand interaction that is yet to be fully understood is exactly how one ligand can act as an agonist in one tissue, and an antagonist in other tissues. The existence of two subtypes of ER, αER and αERx, which reside in different tissues helps to explain some of this selectivity, but the interactions are far more complex. Each ligand that binds to the ER induces a unique conformational change. Different peptide binding sites are exposed in the different complexes; these newly exposed sites are different from other ligand-ER complexes, and different from the native ER as well. This information suggests that each protein-ligand complex is involved in other protein-protein interactions that ultimately affect whether the bound ligand acts as an agonist or antagonist. The distribution of these binding cofactors among tissues may be the reason for one ligand's variable binding profile.[1]

In 1997, Grese et al.[13] published data on 76 raloxifene analogs that they synthesized, including experimental relative binding affinities for each analog, compared to estradiol, as well as IC_{50} values. In the interest of creating a molecule that shares raloxifene's agonist/antagonist profile, but not its susceptibility to degradation, we have used computational methods to study the binding affinity of several of the Grese raloxifene derivatives that have been tested for anti-cancer properties (Figure 4). Each analog was experimentally tested in two assays, a radioligand competitive binding assay as well as an MCF-7 cell proliferation assay, which is used to determine the ligand concentration necessary to inhibit 50% of MCF-7 breast cancer cells (IC_{50}).[13] These IC_{50} values tell us which ligands are good cancer drugs, but the RBA values obtained from the radioligand binding assay will give us a comparison point for the theoretical binding energies.

One of the most difficult aspects of computational modeling of drug-protein interactions is incorporating protein flexibility. Many programs try to model the conformational flexibility of the ligand, while leaving the protein rigid. However, there are several ways to model flexibility in a receptor molecule. One way, the "relaxed complex" method, involves running a long molecular dynamics (MD) simulation of the uncoupled receptor in order to sample the protein's unique conformations. The second phase of this method then involves rapid docking of the ligands into all of the unique protein conformers. The complexes can then be overlaid to evaluate the binding pocket in several modes.[26]

Studies have also shown that for the use of Quantum Mechanical/Molecular Mechanical approaches to study protein-substrate interactions, protein conformational sampling is especially important since small fluctuations in protein shape can have a great effect on the total energy. It has been suggested that the best approach with energy calculations would be to average the results of several protein conformers obtained from a long MD simulation.[27]

The main principle behind computational drug design is that regardless of how specific a particular receptor is for its substrate, we can still create a small molecule that has a greater binding affinity. This kind of assumption has moved away from the historic lock-and-key and induced-fit models. Current understanding has shifted to the idea that a protein exists in an equilibrium of conformational states. It has been suggested in other literature that in order for a ligand to have overall favorable binding affinity, the ligand may be a moderate binder for a highly populated protein conformer, or an excellent binder for a less populated conformer.[28,29]



9j	5,7-CH_3, 6-OH	4'-OH
5b'	6-OCH_3	4'-OH
23c	6-OH	4'-COOH
16b	6-OH	4'-N
5b	6-OH	4'-OCH3
23j	6-OH	4'-CHCH2
11p	6-OH	4'-OH, 3'-CH3
ralox	6-OH	4'-OH
11d	6-OH	2'-CH3

***Figure 4**: Molecular structures of the raloxifene derivatives synthesized by Grese et al. used in our computational modeling.*

Our goal in this study was to benchmark a methodology for computational modeling of the αER-ligand system. The analogs created by Grese et al.[13] were used to explore the binding site of αER. We used AMBER8[30] to study the molecular dynamics of the ER complexes with available crystal structures (ER-estradiol, ER-tamoxifen, ER-raloxifene, shown in Figure 5), as well as αER with the raloxifene analogs. These simulations provided us with information about the conformational flexibility of the protein, as well as the conformational freedom of the ligand. The relative stability of the protein in its entirety encouraged us to focus our energies on the active site of the protein. While it is feasible to run MD simulations on the whole protein-ligand complex, the calculations are very expensive. We created a simpler model of our protein by cutting away all residues that were outside of a 10 Å radius centered on the ligand, and then capped the end groups and held them fixed in order to preserve the active site configuration during the MD simulations. This reduced the size of our system from 4000 atoms down to 1400 atoms. We report the problems we encountered using both methods, and suggest areas that require careful attention in a drug design project such as this one.

***Figure 5**: Top, estradiol (LEFT), raloxifene (purple) and tamoxifen (blue) overlaid in the binding cavity of αER. Right, three crystal structures of ER-estradiol, ER-raloxifene, ER-tamoxifen, overlaid. In each complex, the protein takes on a slightly different conformation.*

Computational Methodology

We compared the crystal structures of estradiol (1ERE),[14] tamoxifen (3ERT)[31] and raloxifene (1ERR),[14] as well as 9 other resolved structures bound to the ER and found that not only was the overall ER protein structure very similar, and that the binding mode for all three, in particular the two SERMs, was nearly identical, as shown in Figure 5. Due to the high resolution of its structure (1.9 Å), the ER from the tamoxifen-ER complex (3ERT) was used as our template. Also, based on crystallographic information of the environment of the ligand, specifically its proximity to an aspartic acid (ASP351), we decided to protonate the amine tail of the ligands for our calculations. We ran MD simulations on each ligand in two systems, the enter ER, and then a smaller model of the ER LBD. To analyze the effect of protonation on the movement and interaction of the amine tail of the ligands, we first used VMD to closely examine the MD simulations. We were able to select all atoms that came within 3 Å of the hydrogen, and from this perspective we could select the atoms of residues that came close enough to hydrogen bond. Once we identified the different residues, we used Ptraj to calculate the distance between atoms for all the possible contacts in each simulation, for both the small system as well as the large system.

Large System

In order to obtain a good starting structure for atomic partial charge assignment, tamoxifen was cut out of 3ERT and underwent a HF/6-31G* geometry optimization using Gaussian03[32] and then was fitted with RESP charges in Antechamber.[33] Hydrogens were added to the ER of 3ERT in the LEaP module of AMBER8.0, and then the entire complex was put through a SANDER minimization using the generalized amber force field (GAFF)[34] for the ligand and ff99SB[35] for the protein. The energy minimization used a standard GB implicit solvation model, 1.2 Å electrostatic scaling, and 2.0 Å van der Waals scaling. We took the minimized complex, removed tamoxifen, and then loaded the empty ER into AutoDockTools.[36] We then used model to dock raloxifene and its derivatives in the ligand binding site. Using the genetic algorithms' default parameters, we docked each ligand 200 times. For each derivative, we chose the conformer that most closely matched raloxifene's crystal structure (1ERR). Each complex was then minimized with SANDER, in the same way as the original. We ran MD simulations on each complex for a duration of 800-3000 ps. The complexes were heated for 50 ps, from 0 K to 300 K, followed by their production runs at a constant pressure. Using Ptraj, in AMBER, the RMS deviation of the ligands during the simulations were calculated in order to determine the point at which the ligand could be considered equilibrated (data not shown). Once that data was obtained, we used the MMGBSA

scoring function in AMBER to calculate changes in free energy of binding for each of the ligands during their equilibrated trajectory.

Small System

We used the program VMD[39] to view the tamoxifen-αER crystal structure, and to select the residues that were within 10 Å of the ligand in the binding pocket. We removed all other residues, except for those that seemed relevant or were connecting the remaining residues, the remaining protein was capped using acetyl and N-methylamine groups. This reduced the size of the protein from 4000 atoms to 1400 atoms (Figure 6). Terminal and "connecting" residues of the cut protein were restrained. The exact ligand pose from the large protein simulation was used for the cut protein system. This was accomplished by direct transfer of coordinates for the ligand to the pdb file of the smaller protein. Each complex was minimized using SANDER, and then put through a heating and a production MD simulation lasting 2000 ps, using the same parameters as the large system, except for the restraints imposed on the selected residues. As with the larger system, Ptraj was used to calculate RMS for the ligands, in order to determine when the ligands were equilibrated (data not shown). To analyze these runs we used both the MMGBSA scoring function, as before, as well as the semi-empirical DivCon[38] single-point calculations (SPCs). We pulled out snapshots from the MD simulations, after the ligand's trajectory was stable, every 50 ps. This structure was then converted to a .in file and submitted to DivCon. We used the PM3 operator, the SCRF solvation method, a nine-angstrom non-bonded cut-off, and the cluster sub-setting (ncore=1 dbuff1=5.5 dbuff2=2.5). The DivCon SPCs were run in triplet in order to get a binding energy ($\Delta H = H_{complex} - (H_{receptor} + H_{ligand})$) for each snapshot. These energies were then averaged to give us a binding energy for a particular ligand.

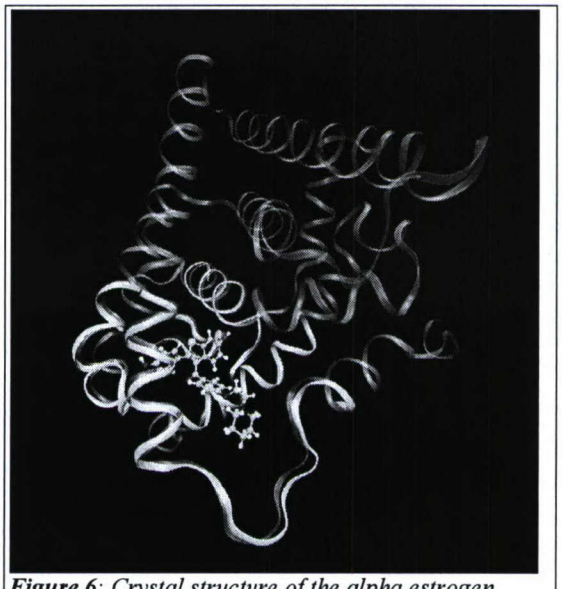


***Figure 6**: Crystal structure of the alpha estrogen receptor with bound inhibitor. Purple indicates full protein, and light blue is the truncated model system.*

Results and Discussion.

The MD simulations for the large protein system showed that the αER LBD remained fairly rigid throughout the production run. However, there was more conformational flexibility seen in the ligand's trajectory, with some trajectories being more stable than others. Using the RMS data, we calculated free energy of binding for the equilibrated ligand-protein complexes. Our initial MMGBSA calculations from the large protein-ligand complex gave us basically no correlation with experimental data, as shown in Figure 7. These poor results, combined with the high computational expense, encouraged us to try and simplify our system.

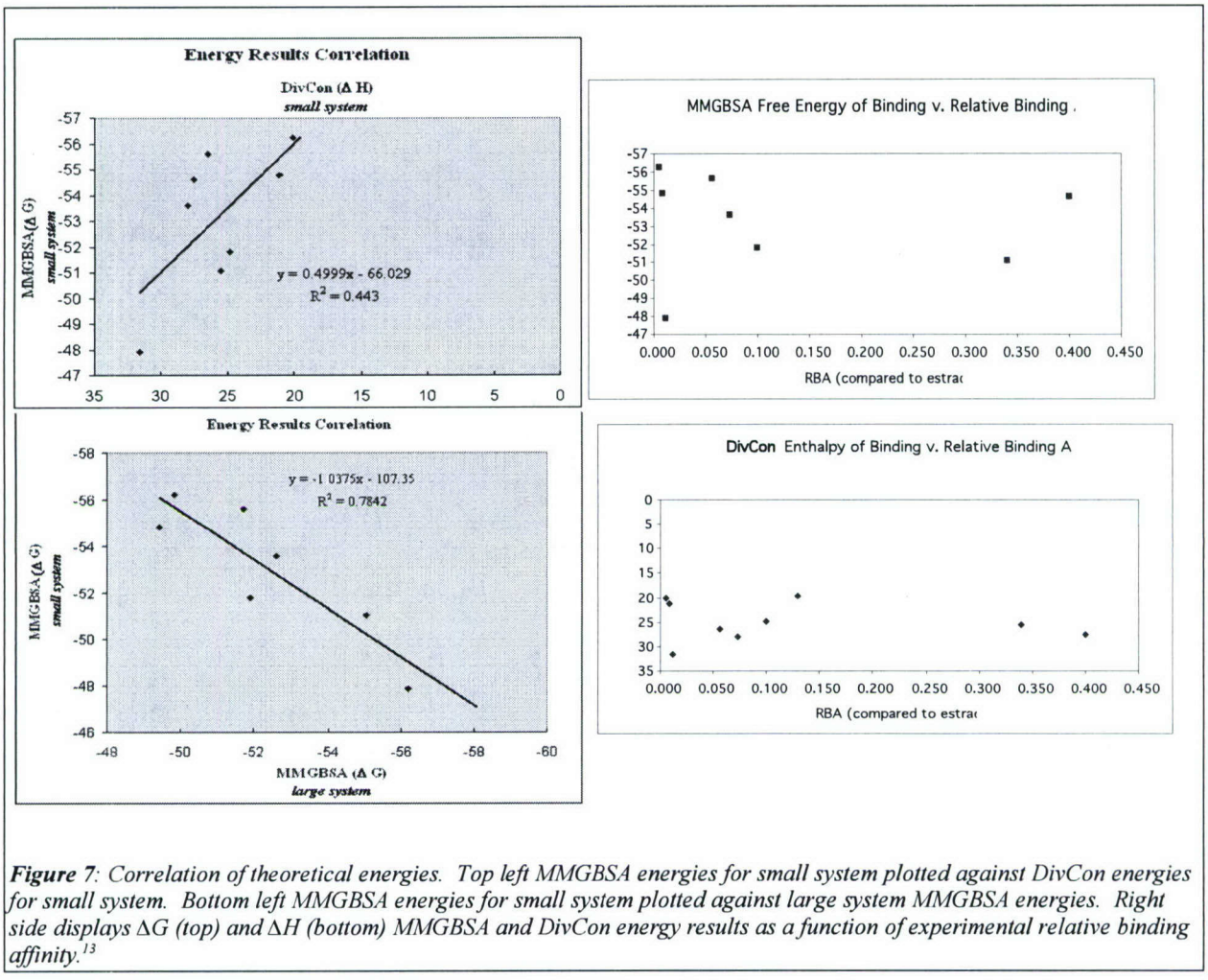


***Figure 7**: Correlation of theoretical energies. Top left MMGBSA energies for small system plotted against DivCon energies for small system. Bottom left MMGBSA energies for small system plotted against large system MMGBSA energies. Right side displays ΔG (top) and ΔH (bottom) MMGBSA and DivCon energy results as a function of experimental relative binding affinity.*[13]

Qualitative observation of the MD simulations showed a significant stability for the majority of the protein. This led us to believe that a significant reduction in the size of the protein would not affect the protein-ligand interaction that we are attempting to model. After cutting down the protein to approximately one third of its original size, and restraining the exterior residues, we ran MD simulations and compared them to the large protein simulations. To understand the qualitative differences, we compared the different raloxifene simulations by superimposing the small system on the larger system in VMD and then watching the simulations progress simultaneously. Figure 8 shows snapshots from this procedure, at the beginning, middle and end. In general, it appeared that the small system held the general shape of the binding cavity, and that raloxifene seemed to have a very similar trajectory in both proteins. Based on this, we assumed that the cut protein system was a reasonable substitution for the larger protein, and we could use these simulations to determine the energetics of ligand-protein interaction. However, when we calculated free energy of binding using first MMGBSA and then change in enthalpy using DivCon, our results, again, did not correlate with the experimental data (Table 1). In order to better understand the discrepancy in our data, we decided to look more closely at what was happening in the MD simulations.

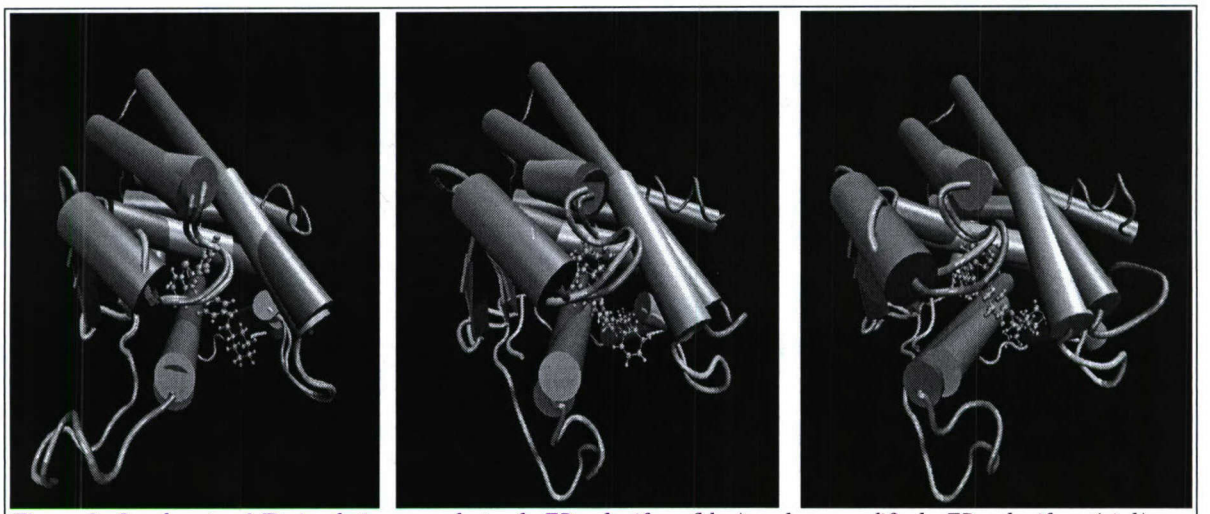


***Figure 8**: Overlapping MD simulation snapshots of αER-raloxifene (blue) and our modified αER-raloxifene (pink). To the far left is the MD simulations at 0 ps, middle at 1000 ps, right at 2000 ps.*

***Table 1**: Experimental relative binding affinity (RBA) data and theoretical energies determined using MMGBSA and DivCon.*

		small system			large system	
	RBA (experimental)	MMGBSA $\Delta G_{kcal/mol}$	(std)	DivCon ΔH (kcal/mol)	MMGBSA $\Delta G_{kcal/mol}$	(std)
9j	0.005	-56.24	(3.4)	20.17	-49.8	(4.0)
5b'	0.008	-54.78	(5.6)	21.12	-49.4	(4.6)
23c	0.012	-47.9	(5.7)	31.58	-56.2	(4.7)
16b	0.056	-55.61	(3.7)	26.51	-51.7	(4.1)
5b	0.073	-53.58	(4.1)	28.02	-52.6	(4.2)
23ja	0.100	-51.79	(3.6)	24.82	-51.9	(4.0)
11p	0.130			19.56	-58.1	(4.7)
ralox	0.340	-51.07	(3.1)	25.54	-55.1	(3.7)
11d	0.400	-54.63	(4.1)	27.54		

The Ptraj RMS data on the ligands (data not shown) allowed us to quantitatively examine the movements of the different ligands. If we examine just the large protein MD runs, we see that the strongest binders, as predicted by RBA, show the most stable trajectories. Raloxifene and tamoxifen show an initial jump in RMS, but stabilize by 200 and 300 ps respectively. The best binder, 11d, shows almost no fluctuation from the very beginning of the simulation. The less potent raloxifene analogs show more varied trajectories suggesting possibly a less tight fit. If we look at these same ligands in the smaller protein system, we see different results. Several of the weak binders show very stable trajectories (9j, 5b', 16b), while raloxifene and 11d seem to have more fluctuation. If we compare the two different simulations for the same ligand, we see some important differences in the strong binders. Raloxifene jumps up to a higher RMS value around 3Å and stays there in the large protein, while in the small protein raloxifene's RMS values are far more varied but most of the time are between 1Å and 2Å. The opposite is true of the experimentally best binder, 11d, whose large protein simulation keeps the ligand at or below 1Å RMS, but in the small system the ligand stays at approximately 2.5Å, but has more fluctuations. The RMS values, by themselves, are not as telling because they are dependent on how good the initial fitted structure is. The part of the ligand that moves the least is the part tucked up into the protein, and consists

primarily of three aromatic rings that are fairly rigid. The part of the ligand that rotates and provides most of the movement is the "tail" that has six rotatable bonds and exposed to bulk solvent. To further understand the MD simulations of these cut protein systems, we looked at the protonated amine group on the tail section of the ligand.

Surprisingly, the protonated amine invariably moves away from ASP351 and hydrogen bonds to other oxygen atoms throughout the binding pocket, as well as one nitrogen and one sulfur (see Figure 9). The protonated amine for the majority of the simulations, forms an *intra*molecular hydrogen bond. The tail of the ligand folds back onto itself and it able to hydrogen bond to an oxygen atom in the tail (see Figure 10). The protonated amine tail has a pK_a of approximately four, which means that when it is isolated in solution, at a physiological pH, the majority of the ligands would not be protonated. This intramolecular hydrogen bond that we see predominate in most of the simulations is likely an artificial result. This is a very difficult issue, since the protonation state of the ligand's tail is likely to depend on dynamics. Because the tail portion of the ligand is solvent accessible, we expect that the amine tail is protonated in the right environment, such as in proximity to ASP351, but unprotonated when completely solvated by bulk water. At best, it is uncertain that the amine portion of the ligand would be protonated in all of the conformations visited by the MD simulations. This result lead us in several different directions looking for an explanation.

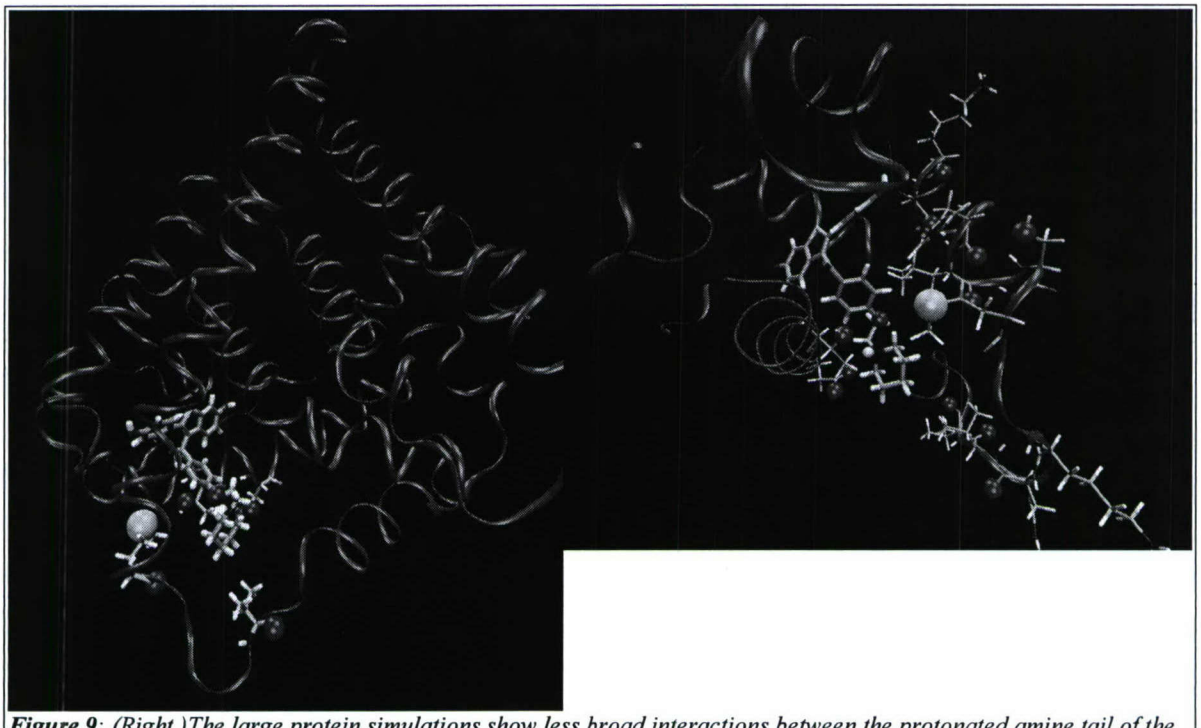


***Figure 9**: (Right)The large protein simulations show less broad interactions between the protonated amine tail of the ligand and the binding cavity residues when compared to the (Left) small protein simulations. The atoms on the residues that interact with the proton*

In some of the simulations, the initial ligand position obtained using AutoDock Tools did not place the hydrogen atom close enough to the aspartic acid to be within hydrogen bonding distance. We assumed that the MD simulation would form the hydrogen bond over the course of the simulation, but this was not the case. However, even in the simulations that began with that hydrogen bond intact, the ligand ultimately changed conformation. We first tried to ascertain whether or not the binding mode includes the hydrogen bond to ASP351, as is suggested in the 1ERR crystal paper.[14] We searched the Protein DataBank for all crystal structures with the αER LBD bound to a ligand. We selected 12 crystals with ligands similar to raloxifene or estrogen and overlaid them in VMD. Figure 11 shows the remarkably close conformation adopted by all of the ligands, even with all of the molecular differences. This seems to provide additional support for that particular binding mode, with the hydrogen bond

between ASP351 and a proton donor on the ligand. There are two possibilities for why our simulations disrupt this bond: the force fields that we are using are inadequately representing the energetics of the hydrogen bond, or the crystal structure is only accurate for the solid phase complex, and is unrealistic at room temperature in solution. We plan to test the second hypothesis by calculating energies based solely on the MD snapshots that maintain the hydrogen bond with ASP351.

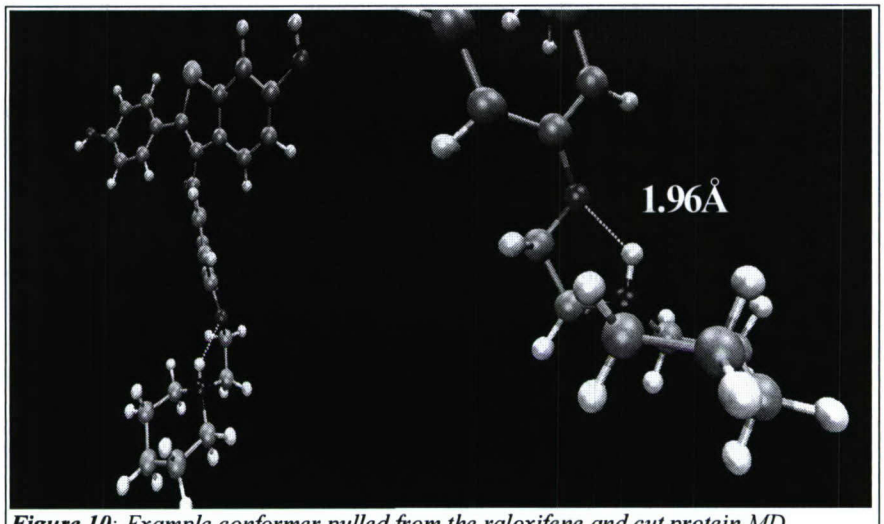


***Figure 10**: Example conformer pulled from the raloxifene and cut protein MD simulation showing how the protonated raloxifene curls back on itself and makes an intramolecular hydrogen bond.*

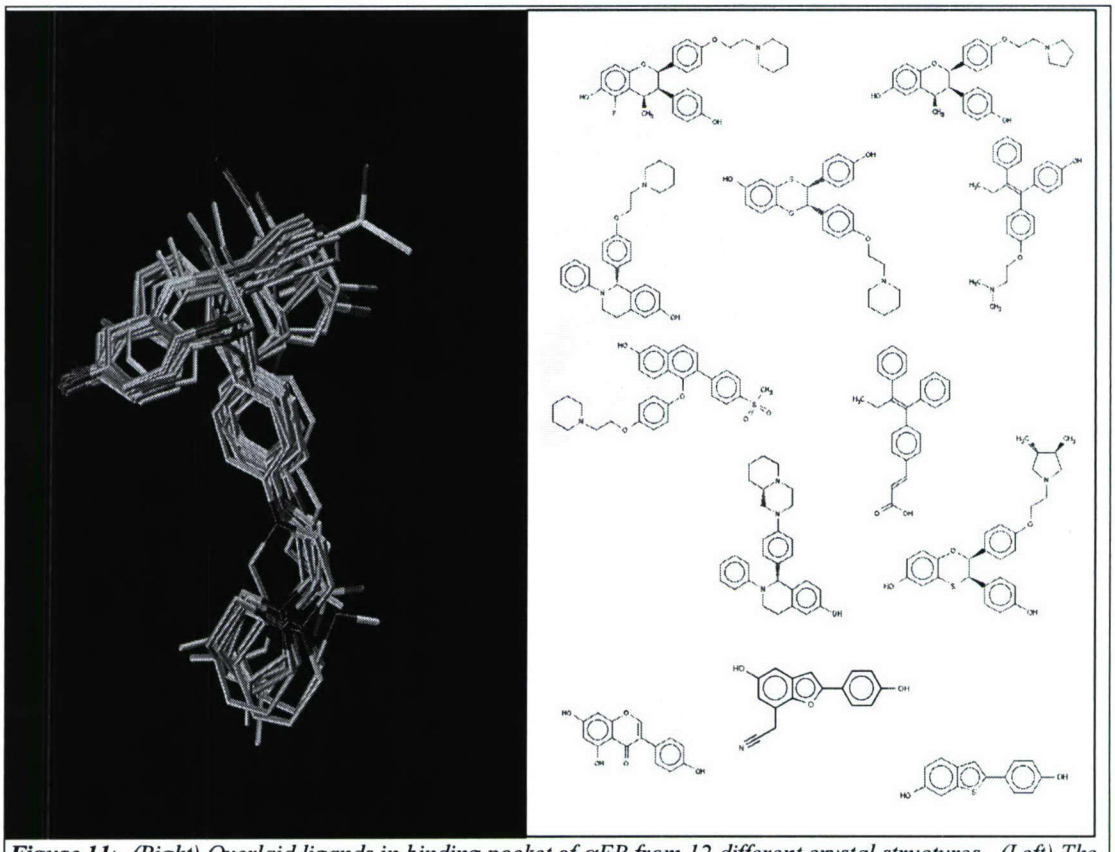


***Figure 11**: (Right) Overlaid ligands in binding pocket of αER from 12 different crystal structures. (Left) The molecular structures of the SERMs/estrogen agonists bound to the αER.*

Since we have very poor correlation to experimental data using two different energy calculations, we compared our theoretical results to each other. In Figure 7 we first show the MMGBSA results plotted against the DivCon results for the small system. The two different methods, and different energy quantities, show reasonable positive correlation. That is to say that both methods agree on the ligands that are strong binders, and those that are poor binders.

Recent Results and Discussion.

Recently, we have obtained positive correlation between our calculated ensemble PM3 energies and the experimental RBA values determined by Grese and coworkers.[13] This was accomplished thanks to two changes in our protocol. First, we used the smaller model protein to approximate the full ER. Second, we immersed the whole system in an octahedral box of TIP4P waters. The force fields and simulations protocols, aside from the changes required for an explicit periodic boundary simulation, are identical to those outlined above. Figure 12 displays the binding energy (ΔH°_{298K}) versus the natural log of the RBA of several ligands, along with the best-fit line and its R^2 value. (If we remove the bottom most right point, our correlation goes above R^2=0.7) The existence of this correlation is very promising, and suggests several things concerning our development of a computational assay. First, that our model of the estrogen receptor (ER) protein is sufficient for modeling the important interactions between the ER and the ligands. Second, our protocols for determining the ligand's partial atomic charges and the docking of these ligands into the binding site are satisfactory for the MD simulations. Third, the simulations are sufficiently sampling the local potential energy surface for obtaining an ensemble of conformations. Finally, that the semiempirical PM3 method as implemented in DivCon can be used to provide binding energies that when averaged over an ensemble of conformations corresponds well to experimental RBA trends.

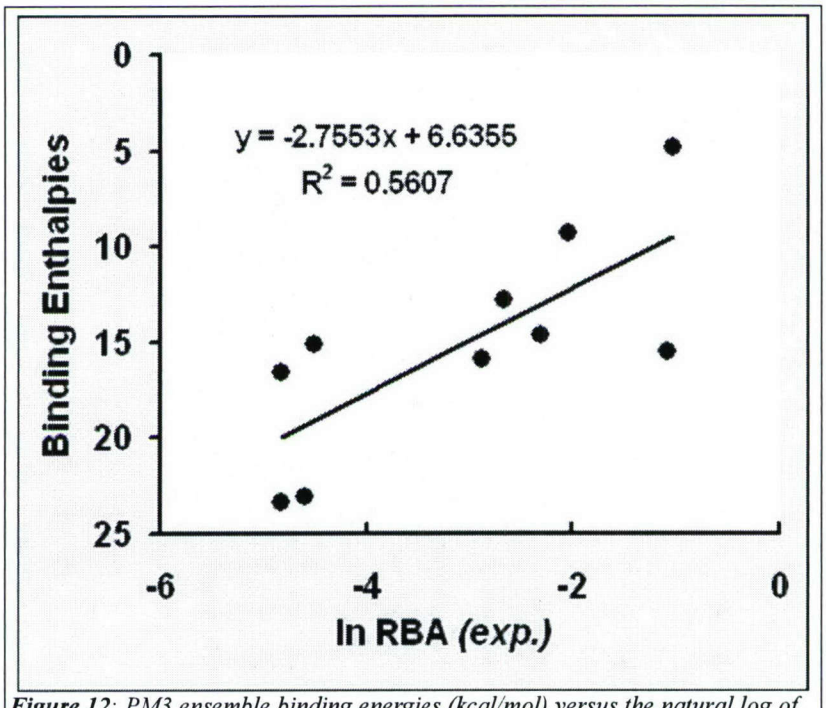


***Figure 12**: PM3 ensemble binding energies (kcal/mol) versus the natural log of the experimental RBAs from several ligands reported in reference 13. The best fit line, its equation, and the R^2 factor are also present.*

Currently we are continuing our analysis of each octrahedral box MD simulation in order to understand the details of each of the ligands interactions with the protein. We hope to be able to isolate specific interactions that can be used to create promising molecular leads for drug discovery. We also hope to improve the correlation by including more ligands, and by enlarging the ensemble that is used in

determining the binding energy for each point. We hope to obtain an R^2 value above 0.8.

***Figure 13**: (Left) Octahedral box surrounding the small ER model of the protein. (Right) Detailed view of the ligand in the ER binding domain with important mediating waters distinguished in gold.*

A few interactions are readily identified as important for stabilizing the interaction between the protein and the ligand. A few important water molecules bridge the ligands to the protein. The presence of waters around arginine and glutamic residues in the binding site helps to maintain a very strong hydrogen bond interaction between glutamic acid and the ligands' 6-OH. (Figure 13) Normally, the arginine and the glutamic acids formed a strong hydrogen bond in our previous implicit solvations. During these implicit solvation simulations, the hydrogen bond between the ligands' 6-OH group and the glutamic acid is initially present, but is quickly lost for the remainder of the situation. In these explicit simulations the waters disrupt this amino acid-amino acid hydrogen bond, allowing the glutamic acid-ligand hydrogen bond to be long lived. Another interesting fact is that the ligand tail remains completely solvated during these explicit water solvations. In the implicit solvations simulations, the tail of the ligand had great mobility and made hydrogen bond contacts with the protein in several locations. In the explicit solvation the tail gains stability in its motion, which appears to help our determination of the ensemble averaged enthalpy of bindings.

In summary, we have created a computational assay for the binding of ligands to the estrogen receptor. Figure 14 displays our current methodology for running a ligand through our computational assay.

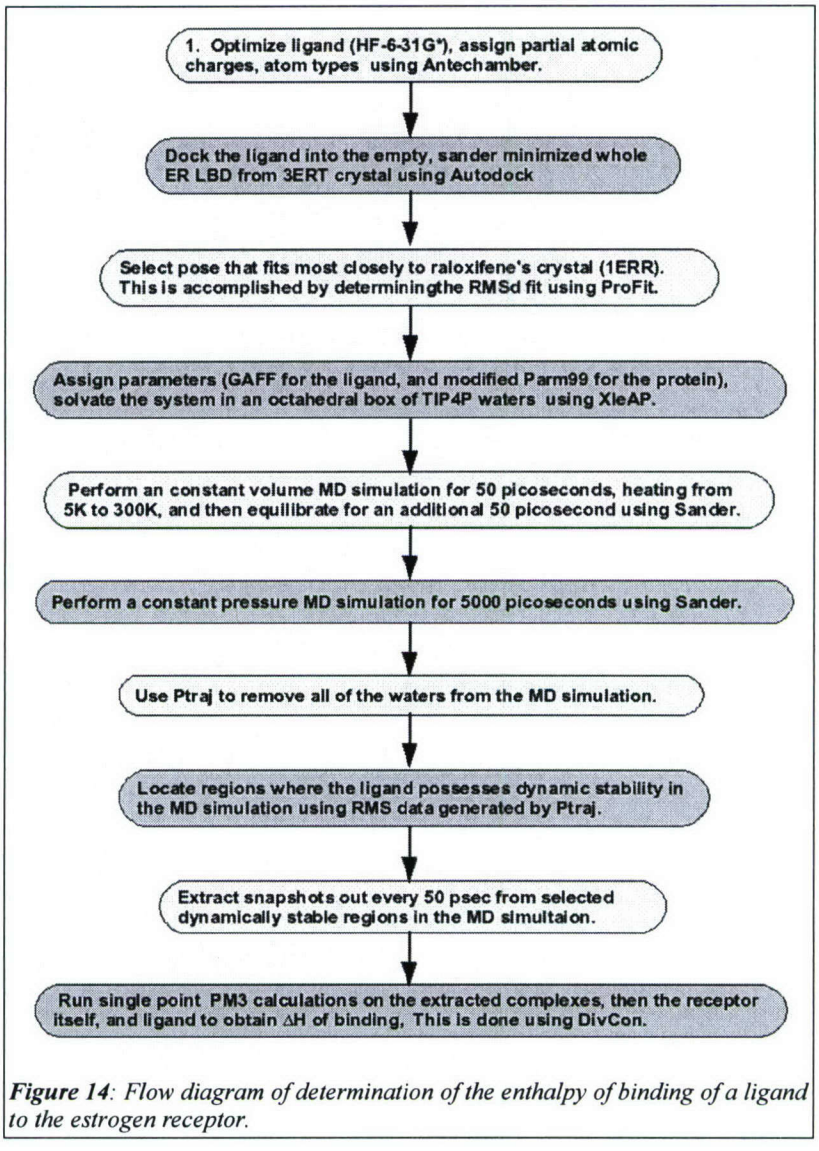


***Figure 14**: Flow diagram of determination of the enthalpy of binding of a ligand to the estrogen receptor.*

Key Research Accomplishments

- Development of a computational model of the Estrogen Receptor binding domain.
- Performed several long molecular dynamics simulations (800 ps - 3 ns) on the ligand-ER system using ligands with known relative binding affinities.
- Determined the optimal method for correlating binding energies to experimental relative binding affinities.
- Creation of a computational assay for ligand binding that involves combining molecular dynamics and quantum mechanics

Reportable Outcomes

1. "Enediyne Bergman Cyclization of the Chemically Active Region of Esperamicin A1" Frank C. Pickard IV,* Chantelle Rein,* Steven Feldgus, Karl N. Kirschner and George C. Shields, in preparation for Journal of Physical Chemistry B .
2. "Computational Design and Experimental Discovery of an Anti-estrogenic Peptide Derived from Alpha-Fetoprotein" Karl N. Kirschner, Katrina W. Lexa,* Amanda M. Salisburg,* Katherine A. Alser,* Leroy Joseph, Thomas T. Andersen, James A. Bennett, Herbert I. Jacobsen, and George C. Shields, in preparation for Nature.
3. "The Nature of the Dissociation of H_2O in Water" Sarah T. Tschampel,* Frank C. Pickard IV,* Jennifer R. Derby,* and George C. Shields, in preparation.
4. "In Search of $CS_2(H_2O)_n$, n=1-4 Clusters" Gregory M. Hartt,* Timothy M. Evans,* Karl N. Kirschner and George C. Shields, J. Chem. Phys. submitted 9/6/06, in revision.
5. "Do Hydroxyl Radical-Water Clusters, $OH(H_2O)_n$, n=1-5, Exist in the Atmosphere?" Marco A. Allodi,* Meghan E. Dunn,* Jovan Livada,* Karl N. Kirschner and George C. Shields, J. Phys. Chem. A accepted for publication 9/28/2006.
6. "Comparison of CCSD(T), W1, and other Model Chemistry Predictions for Gas Phase Deprotonation Reactions" Frank C. Pickard IV,* Daniel R. Griffith,* Skylar J. Ferrara,* Matthew D. Liptak,* Karl N. Kirschner and George C. Shields, Int. J. Quantum Chem. 106 (2006) 3122-3128.
7. "Exploration of the Potential Energy Surfaces, Prediction of Atmospheric Concentrations, and Vibrational Spectra of the HO_2***$(H_2O)_n$ (n=1-2) Hydrogen Bonded Complexes" Kristin S. Alongi,* Theodore S. Dibble, George C. Shields and Karl N. Kirschner, J. Phys. Chem. A 110 (2006) 3686-3691.
8. "Computational design of a small peptide that inhibits breast cancer: An overview of computational chemistry research at Hamilton College." 38th ACS Middle Atlantic Regional Meeting, June 5, 2006. Presented by George C. Shields
9. "Analysis of the Affinity of Raloxifene and its derivatives to the α-Estrogen Receptor Ligand Binding Domain," Karilyn Larkin,* Amy Barrows,* Karl Kirschner and George Shields at the 46th annual Sanibel International Symposium on Atomic, Molecular, Biophysical, and Condensed Matter Theory, March 2006, St. Simons, GA. Presented by K. Larkin and won the Outstanding Poster Presentation Award.
10. "Generalized Born Molecular Dynamics and MMPBSA Analysis of the α-Estrogen Receptor bound to Raloxifene, Tamoxifen and Estradiol," Amy Barrows,* Karilyn Larkin,* Karl Kirschner, and George C. Shields at the Fourth MERCURY Conference in Computational Chemistry, July 2005, Clinton, NY. Poster presented by A. Barrows.

11. "Analysis of the Affinity of Raloxifene and its derivatives to the α-Estrogen Receptor Ligand Binding Domain," Karilyn Larkin,* Amy Barrows,* Karl Kirschner and George Shields at the Fourth MERCURY Conference in Computational Chemistry, July 2005, Clinton, NY. Presented by K. Larkin.
12. Karilyn Larkin – Chemistry Thesis: "Analysis of the Affinity of Raloxifene and its Derivatives to the α-Estrogen Receptor Ligand Binding Domain: A Model for Drug Design" 2005-2006. Currently pursuing an M.D., Albany Medical School, NY.

Conclusion

We have successfully combined MD & QM methods in an innovative way to develop a computational assay that will allow for the development of novel lead compounds that should be active against breast cancer. We have determined the optimal method for correlating binding energies to experimental relative binding affinitites (RBAs). This assay will be used to design novel inhibitors with high-predicted activity. Examining the dynamics of the ligands that exhibit high affinity versus those that do not, we will be able to determine the important ligand-protein interactions. This will allow for the specific modifications to be made to this ligands in hopes of enhancing the binding affinities. The new ligands will be tested in our computational assay, and the promising ones will be synthesized and tested in experimental ER assays.

Future Directions

We will continue to refine our assay, by including more MD data points in our determination of the enthalpy of binding. We will also investigate the protonation state of the amine on the tail of our analogues. Now that we observe that the tail stays solvated during the simulation, we believe that the amine will exist in the unprotonated state. It is unlikely that this will alter the correlation.

We will perform our computational protocol on new ligands from the Renaud paper[39] for protonated and unprotonated runs. Future work will involve a collaboration with medicinal chemists who will synthesize the new molecules, test them in an ER assay, and evaluate the most promising compounds against the MCF-7 breast cancer cell line.

References:

1. Paige L, Christensen D, Gron H, Norris J, Gottlin E, Padilla K, Chang C, Ballas L, Hamilton P, McDonnell D, Fowlkes D. Estrogen receptor (ER) modulators each induce distinct conformational changes in ER α and ER β. *Proceedings of the National Academy of the Sciences of the United States of America* **96** (1999) 3999-4004.

2. Schmidt, J M, Mercure J, Tremblay GB, Page M, Feher M, Dunn-Dufault R, Peter MG, Redden PR. De novo design, synthesis and evaluation of a non-steroidal diphenylnaphthyl propylene ligand for the estrogen receptor. *Bioorganic and Medicinal Chemistry* **11** (2003) 1389-1396.

3. Klinge C.M., Jernigan S.C., Mattingly K.A., Risinger K.E., Zhang J. Estrogen response element-dependent regulation of transcriptional activation of estrogen receptors α and β by coactivators and corepressors. *Journal of Molecular Endocrinology* **33** (2004), 287-410.

4. Klinge CM, Kaur K, Swanson HI. The aryl hydrocarbon receptor interacts with estrogen receptor (AHR)/AHR nuclear translocator (ARNT) heterodimers interacts with naturally occurring estrogen response elements. *Molecular and Cellular Endocrinology*. **157** (1999) 105-119.

5. Meegan, M.J., Lloyd, D.G. Advances in the science of estrogen receptor modulation. *Current Medicinal Chemistry* **10** (2003) 181-210.

6. Kuiper GGJM, Enmark E, Pelto-Huikko M, Nilsson S, Gustafsson JA. Cloning of a novel estrogen receptor expressed in rat prostate and ovary. *Proceedings of in National Academy of Sciences USA* **93** (1996) 5925-5930.

7. Pike MC, Spicer DV, Dahmoush, Press MF. Estrogens, progestogens, normal breast cell proliferation, and breast cancer risk. *Epidemiology* Review **15** (1992) 17-37.

8. Holmes FA, Liticker J. Pharmacogenomics of tamoxifen in a nutshell – and who broke the nutcracker? *Journal of Oncology Practice* **1** (2005) 155-159.

9. Ali S, Coombes RC. Estrogen receptor alpha in human breast cancer: occurrence and significance. *Journal of Mammary Gland Biology and Neoplasia.* **5** (2000) 271-281.

10. Tobias JH, Compson J. Does estrogen stimulate osteoblast function in postmenopausal women? *Bone* **24** (1999) 121-124.

11. Dhandapani K, Brann D. Protective Effects of Estrogen and Selective Estrogen Receptor Modulators in the Brain. *Biology of Reproduction* **67** (2002) 1379-1385.

12. Curtis RE, Freedman MD, Sherman ME, Fraumeni Jr JFR. Risk of malignant mixed mullerian tumors after tamoxifen therapy for breast cancer, *Journal of the National. Cancer Institute* **96** (2004) 70–74.

13. Grese TA, Cho S, Finley DR, Godfrey AG, Jones CD, et. al. Structure-activity relationships of selective estrogen receptor molecules: modifications to the 2-arylbenzothiophene core of raloxifene. Journal of Medicinal Chemistry **40** (1997) 146-167.

14. Brozowski AM, Pike CW, Dauter Z, Hubbard RE, Bonn T, Engstrom O, Ohman L, Green GL, Gustafsson JA, Carlquist M. Molecular basis of agonism and antagonism in the oestrogen receptor. *Nature* **389** (1997) 753-758.

15. Genant HK, Lang T, Fuerst T, Pinette KV, Zhou C, Thiebaud D, Diez-Perez A. Treatment with raloxifene for 2 years increases vertebral bone mineral density as measured by volumetric quantitative computed tomography. *Bone* **35** (2004) 1164-1168.

16. Cox D, Sarkar S, Harper K, Barrett-Connor E. Effect of Raloxifene on the incidence of elevated low density lipoprotein (LDL) and achievement of LDL target goals in postmenopausal women. *Current Medical Research and Opinion* **20** (2004) 1049-1055.

17. Cummings S, Eckert S, Krueger K, Grady D, Powles T, Cauley J, et al. The effect of raloxifene on risk of breast cancer in postmenopausal women. *Journal of the American Medical Association* **281** (1999) 2189-97.

18. Cauley J, Norton L, Lippman ME, Eckert S, Krueger K, Purdie DW, et al. Continued breast cancer risk reduction in postmenopausal women treated with raloxifene: 4-year results from the MORE trial. *Breast Cancer Research and Treatment* 65 (2001) 125-134.

19. Martino S, Cauley J, Barrett-Connor E, Powles T, Mershon J, Disch D, Secrest R, Cummings S. Continuing Outcomes Relevant to Evista: breast cancer incidence in postmenopausal osteoporotic women in a randomized trial of raloxifene. *Journal of the National Cancer Institute* **96** (2004) 1751-1761.

20. Hulley S, Grady D, Bush T, et al. Randomized trial of estrogen plus progestin for secondary prevention of coronary heart disease in postmenopausal women. Heart and Estrogen/progestin Replacement Study (HERS) Research Group. *Journal of American Medical Association* **280** (2001) 1243-9.

21. Keech C, Sashegyi A, Barrett-Connor E. Year-by-year analysis of cardiovascular events in the Multiple Outcomes of Raloxifene Evaluation (MORE) trial. *Current Medical Research and Opinions* **21** (2005) 135-140.

22. Kemp D, Fan P, Stevens J. Characterization of Raloxifene Glucuronidation in Vitro: Contribution of Intestinal Metabolism to Presystemic Clearance. *The American Society for Pharmacological and Experimental Therapeutics* **30** (2002) 694-700.

23. Hochner-Celnikier D. Pharmokinetics of raloxifene and its clinical application. *European Journal of Obstetrics and Gynecological Reproductive Biology* **85** (1999) 23-29.

24. Jeong E, Liu Y, Lin H, Hu M. Species- and Disposition Model-Dependent Metabolism of Raloxifene in Gut and Liver: Role of UGT1A10. *Drug Metabolism and Disposition* **33** (2005) 785-794.

25. Jeong E, Lin H, Hu M. Disposition Mechanisms of Raloxifene in the Human Intestinal Caco-2 Model. *Journal of Pharmacology and Experimental Therapeutics* **310** (2004) 376-385.

26. Lin J, Perryman A, Schames J, McCammon J. Computational Drug Design Accommodating Receptor Flexibility: The Relaxed Complex Scheme. *Journal of the American Chemical Society* **124** (2002) 5632-5633.

27. Klahn M, Braun-Sand S, Rosta E, Warshel A. On Possible Pitfalls in ab Initio Quantum Mechanics/Molecular Mechanics Minimization Approaches for Studies of Enzymatic Reactions. *Journal of Physical Chemistry* **109** (2005) 15645-15650.

28. Carlson H. Protein flexibility and drug design: how to hit a moving target. *Current Opinion in Chemical Biology* 6 (2002) 447-452.

29. Carlson H, McCammon A. Accommodating Protein Flexibility in Computational Drug Design. *The American Society for Pharmocological and Experimental Therapeutics* **57** (2000) 213-218.

30. D.A. Case, D.A. Pearlman, J.W. Caldwell, T.E. Cheatham III, J. Wang, W.S. Ross, C.L. Simmerling, T.A. Darden, K.M. Merz, R.V. Stanton, A.L. Cheng, J.J. Vincent, M. Crowley, V. Tsui, H. Gohlke, R.J. Radmer, Y. Duan, J. Pitera, I. Massova, G.L. Seibel, U.C. Singh, P.K. Weiner and P.A. Kollman (2002), AMBER 8, University of California, San Francisco.

31. 3ERT Shiau, A.K., Barstad, D., Loria, P.M., Cheng, L., Kushner, P.J., Agard, D.A., Greene, G.L. The structural basis of estrogen receptor/coactivator recognition and the antagonism of this interaction by tamoxifen. *Cell* **95** (1998) 927-937.

32. Gaussian 03, Revision B.02, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople, Gaussian, Inc., Pittsburgh PA, 2003.

33. Ante Wang, J., Wang, W., Kollman P. A.; Case, D. A. "Automatic atom type and bond type perception in molecular mechanical calculations". Journal of Molecular Graphics and Modelling, **25** (2006) 247-260.

34. Gaff Wang, J., Wolf, R. M.; Caldwell, J. W.;Kollman, P. A.; Case, D. A. "Development and testing of a general AMBER force field". Journal of Computational Chemistry **25** (2004) 1157-1174.

35. Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, Simmerly C. (2006) To be published.

37. Morris GM, Goodsell DS, Olsen AJ. AutoDock 3.05. The Scripps Research Institute (1991-2000).

37. Humphrey W, Dalke A, Schulten K. VMD- visual molecular dynamics. *Jounral of Molecular Graphics* **14** (1996) 33-38.

38. Wang B, Raha K, Liao N, Peters MB, Kim H, Westerhoff LM, Wollacott AM, van der Vaart A, Gogonea, Suarez D, Dixon SL, Vincent JJ, Brothers EN and Merz KM Jr., DivCon 4.x, distributed by QuantumBio Inc., State College, PA 16803, (2005).

39. Renaud, J., Bischoff, S.F., Buhl, T., Floersheim, P., Fournier, B., Halleux, C., Kallen, J., Keller, H., Schlaeppi, J.-M., Stark, W. "Estrogen Receptor Modulators: Identification and Structure-Activity Relationships of Potent Eralpha-Selective Tetrahydroisoquinoline Ligands." Journal of Medicinal Chemistry **46** (2003) 2945.

George C. Shields

Curriculum Vitae

Hamilton College
Department of Chemistry
198 College Hill Road, Clinton, NY 13323
315-859-4728
gshields@hamilton.edu
http://www.chem.hamilton.edu/faculty/shields.html

EDUCATION

1986-1989	Postdoctoral Research, Yale University and the Howard Hughes Medical Institute; conducted research on Protein-DNA interactions, using X-ray crystallography in the laboratory of Professor Thomas A. Steitz.
1986	Ph.D. in Physical Chemistry, Georgia Institute of Technology, Dissertation Title: "Investigation of Structures and Reactivities of Hydrocarbon Ions Through Gaseous Charge Transfer Reactions" Thesis advisor, Professor Thomas F. Moran
1983	MS in Chemistry, Georgia Institute of Technology; Thesis Title: "Charge Transfer Cross Sections for Collisions of Ar^{2+} Ions with Various Target Gases"
1981	BS in Chemistry, Georgia Institute of Technology, Undergraduate research project in organic synthesis, with Professor Charles L. Liotta

EMPLOYMENT HISTORY

2003-	Winslow Professor of Chemistry, Hamilton College
2005-	Director, Center for Molecular Design
1999-2006	Chair, Department of Chemistry, Hamilton College
1998-	Professor, Department of Chemistry, Hamilton College
1996.1998	Chairperson, Department of Chemistry, Lake Forest College
1995-1996	Invited Professor, University of Barcelona, Department of Biochemistry and Molecular Biology
1994-1998	Associate Professor of Chemistry, Lake Forest College
1989-1994	Assistant Professor of Chemistry, Lake Forest College; taught Physical Chemistry, Biophysical Chemistry, General Chemistry, and directed undergraduate research students
1986.1989	Postdoctoral Associate, Yale University & Howard Hughes Medical Institute, with T.A. Steitz
1980-1985	Teaching Assistant for General Chemistry and Physical Chemistry Laboratories at the Georgia Institute of Technology

EXTERNAL GRANT SUPPORT

NSF-MRI 2005-2008 "Acquisition of a Linux Cluster for the Molecular Education and Research Consortium in Undergraduate computational chemistRY (MERCURY)"

NIH-AREA 2005-2008 "Reactivity and Dynamics of Enediyne Natural Products"

DOD-Concept 2005-2006 "Development of a Computational Assay for the Estrogen Receptor"

NSF-RUI 2005-2008 "RUI: Calculating Acid Dissociation Constants in Aqueous Solution"

Research Corporation 2004-2006 "Quantum Chemical Investigation of the Mechanism of Action of the Enediyne Natural Products"

NSF-MRI 2003-2006 "MRI/RUI- Acquisition of a High Field NMR for Chemistry Research" (co-PI along with I. Rosenstein & H. Lehman; R.B. Kinnel PI)

Dreyfus-Special Grant Program in the Chemical Sciences, 2003-06 "The Development of a National Model for Increasing the Number of Chemistry Majors"

NSF-STEP 2002-2007 "STEP: A Multidisciplinary Bridge Program in a Liberal Arts Setting"

ACS/PRF 1/1/02-8/31/05 "Accurate Calculation of pK_a Values"

NYS Health Research Science's Board-EMPIRE, 1/1/02-12/31/03 "Design of Molecules that Inhibit Human Breast Cancer"

NSF-MRI 9/1/01-8/31/04 "Acquisition of High Performance Computers for the Northeastern Undergraduate Research Chemistry Consortium"

MERCK/AAAS Undergraduate Research Program, 7/1/00-6/30/03, (w/ J. Garrett, Biology) "Merck/AAAS Hamilton Summer Research Scholars Program"

Camille and Henry Dreyfus Scholar/Fellow Program, 7/1/00-9/30/02, "Quantum Biology & Computational Biochemistry"

NIH-AREA, 6/15/99-7/14/02, "Computational Design of Haptens to Destroy Cocaine"

American Chemical Society/Petroleum Research Fund 7/1/98-8/31/00 "Catalytic Antibodies: Correlation of Hapten Structure with Induction of Catalysis"

North Atlantic Treaty Organization 1/1/98-12/31/99,, Collaborative Research Grant between the NATO countries Spain, England, and the United States. "Studies on DNA (t) Structure and Reactivity"

Spanish Ministry of Science and Education, 6/1/95-5/30/96, , Sabbatical Research Award

National Science Foundation 1994-1996 "Development of Computational Chemistry and Computer Graphics Laboratory Experiments and Manuals"

NIH-AREA 1992-1995 "Cloning and Crystallization of Isoschizomers of Pst I"

Research Corporation 1990-1992 "Crystallization & Characterization of Circumsporozoite Proteins of the Malaria Parasite *P. falciparum*"

Since 1990 I have received - in external grant support for 20 proposals funded by NSF, NIH, ACS/PRF, Dreyfus Foundation, Research Corporation, and other funding agencies.

HONORS AND AWARDS

2005	Editorial Board of Computing Letters
2003	Winslow Chair in Modern Science
2000-02	Camille and Henry Dreyfus Scholar
1995-96	Spanish Ministry of Education and Science Sabbatical Award
1993	William L. Dunn Award for Outstanding Teaching and Scholarly Promise, awarded by an elected committee of faculty at LFC
1985	Conoco/Du Pont Fellowship, awarded by GT Chemistry Faculty
1984	Elected full member of Sigma Xi Research Society
1984	Sigma Xi Outstanding Master's Thesis Award in Science
1983	Omicron Delta Kappa (ODK), National Leadership Honorary Society
1980	Coaches Award in Lacrosse, for Leadership On and Off the Field
1978	Full Four Year Tuition Baker Scholarship Recipient
1977	New York State Regents Scholarship

UNDERGRADUATE TEACHING

Hamilton College

Physical Chemistry; Physical Chemistry Laboratory; Principles of Chemistry; Biological Chemistry; Research.

Lake Forest College

Physical Chemistry 1989-1998; Physical Chemistry Laboratory 1989-1998; Principles of Chemistry 1991-1998; Topics in Biochemistry 1990-1998; Research & Senior Thesis in Chemistry 1989-1998; General Chemistry Laboratory 1989-1990, 1993; Principles of Chemistry Project Laboratory 1991-1993; Co-director, Richter Apprentice Scholar Summer Program 1992, 1994.

NATIONAL SPONSORED STUDENT AWARDS - Since 1995 my students have received twelve national awards, including two Fulbrights and six Goldwaters.

Barry M. Goldwater Scholarship, Daniel Griffith, 2006-07,
Barry M. Goldwater Scholarship, Mary Beth Day, 2005-07,
Barry M. Goldwater Scholarship, Meghan Dunn, 2005-06,
Fulbright Grant, Spain, Sarah Taylor, 2003-04, living expenses for a year in Barcelona
Pfizer Summer Undergraduate Research Fellow, Matthew Liptak, 2002,
Barry M. Goldwater Scholarship, Matthew Liptak, 2002-03,
Barry M. Goldwater Scholarship, Lorena Hernandez, 2001-03,
Barry M. Goldwater Scholarship, Ann Marie Toth, 2001-02,
ACS Scholars, Lorena Hernandez, 1999-2002,
Gates Millennium Scholar, Lorena Hernandez, 2000-01,
Fulbright Grant, England, Edward Sherer, 1997-98, living expenses for a year in Nottingham
CUR Summer Research Fellowship, Gordon Turner, 1995, (declined)

UNDERGRADUATE RESEARCH DIRECTION – *Since 1989 I have supervised a total of 110 undergraduates (58 F) and 3 high school students in research projects. Of 49 graduates (27F), 21 have entered PhD programs (10F); two entered Law School (2F); thirteen entered professional schools for health professions (9F), seven entered MS or MAT programs (4F); and three are secondary school teachers (F).*

Hamilton College - *supervised 92 undergraduates (49 F) & 3 secondary school students*

Graduates *Of 31 HC graduates (18 F), 11 entered Ph.D. programs (6F); two entered Law School (2F); 10 entered professional schools for health professions (7F); two are high school chemistry and physics, or math teachers (F).*

2006

Valery Danilack, Senior Thesis in biochemistry, "Testing the Estrogen Receptor Binding Affinity of Pharmacophores of Raloxifene in Search of Anti-Breast Cancer Drugs" M.S. Public Health, Yale University

Ashley Deline, Senior Thesis in chemistry, " Molecular Dynamic Simulations of a Glycoform and its Constituent Parts Related to Rheumatoid Arthritis" Fred Hutchingson Cancer Center

Meghan Dunn, Senior Thesis in chemistry, "Hydrated N_2 and O_2 Complexes: Implications of Weakly Bound Species in Atmospheric Chemistry" Ph.D. Chemistry, University of Colorado, Boulder.

Karilyn Larkin, Senior Thesis in chemistry, "Analyses of the Affinity of Raloxifene and its Derivatives to the α-Estrogen Receptor Ligand Binding Domain: A Model for Drug Design" M.D. Albany Medical School

Amber Gillis, Biochemistry, MAT Chemistry & Biology, Union College

Rebecca Shepherd, Biology, work.

2005

Timothy Evans, Senior Thesis in chemistry, "Hydrated Complexes of Atmospheric Sulfur Species" M.S. Mathematics, NYU

JunChan Hong, Chemistry, Analyst at UBS Financial Services

Scott Huntington, Senior Thesis in chemistry, "Computer Aided Design of Selective Estrogen Receptor Modulators", NYU Medical School

David Kelland, Futures Trader, Gelber Group

Katrina Lexa, Senor Thesis in chemistry, "Conformational Analysis of Peptides: Modeling a Breast Cancer Inhibiting Peptide derived from AFP", M.S. Public Health, U. Michigan; Ph.D. Medicinal Chemistry, U. Michigan.

Frank Pickard, Senior Thesis in chemical physics, "The Enediyne Anticancer Antibiotics: A Study of the Bergman Cyclization Energy Barriers of Esperamicin A_1" Research Assistant, Hamilton College; Ph.D. Physical Chemistry, U. Georgia.

Hima Poonati, Chemistry, St. George's Medical School.

Andreu Viader Valls, Biochemistry, Ph.D. Biochemistry, Washington University

2004

Gabrielle Markeson, Senior Thesis in chemistry, "Anti-Breast Cancer Drug Design" Eltron Research in Boulder, CO. Fordham Law School, 2005.

Emma Pokon, minor in Chemistry, Law School, University of Vermont.

Brent Matteson, Chemical Physics, Graduate School at Oregon State University

Sinan Misirli, Senior Thesis in chemistry, "Theoretical pK_a Calculations of Ionic Species: Implicit and Explicit Solvation Methods" Work.

2003

Damien Ellens, Senior Thesis in chemistry, "Conformational Analysis of Cyclic Peptides: Evaluating the relationship between structure and energetics using Molecular and Quantum Mechanics" Yale Medical School.

Lorena Hernandez, biochemistry, Goldwater Scholar, 2001-02; currently MD/PhD student at Albert Einstein College of Medicine.

Matthew Liptak, Senior Thesis in chemistry, "Modeling the Inhibition of Cdc25B: Incorporating QM/MM into Rational Drug Design" Seven publications from his undergraduate research, publications 33, 34, 36, 38, 39, 43 & 48; Goldwater Scholar, 2002-03; currently graduate student in chemistry, University of Wisconsin (Tom Brunold).

Chantelle Rein, Senior Thesis in biochemistry, "An Investigation into the Usefulness of the ONIOM QM/MM Method in Determining the Energetic Pathways of Esperamicin A_1" Graduate school in biochemistry, Oregon Health Sciences University.

Sarah Taylor, Senior Thesis in biochemistry, " Computational Approaches to Anti-Breast Cancer Drug Design" Fulbright Award to Barcelona, Spain (Modesto Orozco, 2003-04). Medical School at the NYU (2004).

2002

Beth Hayes, Senior Thesis in chemistry, "The Mechanisms of DNA Cleavage by Diradicals." One publication from her undergraduate research, publication #45; Currently enrolled at the New York College of Osteopathic Medicine.

Jaime Skiba, Senior Thesis in chemistry, "A Computational Analysis of the AFP-derived Biologically Active Octamer and Some Derivatives." One publication from her undergraduate research, publication #45; High School Chemistry & Physics Teacher at the High School for Environment, Business, and Technology in Williamsburg, NYC.

Annie Toth, biochemistry. Publication #33 from her 1999-2000 research year; Goldwater Scholar, 2001-02; currently graduate student in molecular, cellular and developmental biology at t he University of California, Santa Barbara.

2001

Jennifer Derby, Senior Thesis in chemistry, "The Application of Electrostatic Potential to Drug Design." Currently working as an Applications Specialist at a lab equipment company, Thermo IEC.

2000

Sam Bono, Senior Thesis in chemistry, "The Computational Analysis of the Biologically Active 8-mer Peptide of Alpha-fetoprotein." Publication #35 from his undergraduate research; O.D. Pennsylvania College of Optometry.

Ryan Elias, Senior Thesis in chemistry, "Modelling the Hydroxylase-Phenol Component of Methane Monooxygenase: A Preliminary Study of the Effectiveness of DFT Calculations." Alternate for Fulbright to Italy. Worked for Eziba in NYC. Currently graduate student in food science at the University of Massachusetts, Amherst.

Sarah Tschampel, Senior Thesis in chemistry, "The Computational Design of a Transition State Analog for the Hydrolysis of Cocaine." One publication from her undergraduate research, publication #45; Currently graduate student in chemistry at the University of Georgia (Rob Woods).

James Lempenau, Senior Thesis in biochemistry, "Analysis of Netropsin Binding with A-T Rich DNA" Currently working on Wall Street (J.P. Morgan).

Lake Forest College *supervised 18 students in undergraduate research projects (9F); ten entered Ph.D. programs (4F); four entered MS or MAT programs (2F); three entered professional schools for health professions (2F); and one entered industry (F).*

Graduates

2000

Mariela Serrano, Masters of Biochemical Engineering, University of Delft, The Netherlands.

Becky Turner, completed a Master of Arts in Teaching program at the University of Maine. Currently middle school math and science teacher.

1998

Dale Miller, Honors Senior Thesis "Purification and Crystallization of the poly(dT) ·poly(dA)·poly(dT) DNA Triplex Complexed with the Intercalator, Damian's Salt." Enrolled as a biochemistry graduate student at Colorado State University.

1997

Edward Sherer - Honors Senior Thesis "Catalytic Hydrolysis of Cocaine: Mechanism and Design." Fulbright Graduate Fellowship (England), 1997-98; Five publications from his undergraduate research, publications #25, #26, #28, #29, &#31; Ph.D. completed at the University of Minnesota (Chris Cramer). Currently Research Scientist at Rib-X Pharmaceuticals.

Gordon Turner – Honors Senior Thesis "Computational Design of New Drugs to Treat Cocaine

Addiction." Four publications from his undergraduate research, publications #25, #26, #29, & #31; Ph.D. in physical chemistry completed at Yale University (Charles Schmuttenmaer). Postdoctoral research fellow at Harvard Medical School, in the Department of Radiology, Center for Molecular Imaging Research (Vasilis Nziachristos). Currently employed by Novartis.

1996

Jeff Johnson – Summer research project in X-ray crystallography. Enrolled as a chemistry graduate student at the University of Colorado, Boulder.

Nichole Roberts – Summer research project in X-ray crystallography. M.S. Biology, University of Colorado. Currently in medical school at Boston University.

1995

Heather Brummel - Honors Senior Thesis, recipient of the *Phi Beta Kappa Senior Thesis Award* "Semiempirical Study of the Bergman Reaction: A Computationally Efficient and Accurate Method for Modeling the Enediyne Anticancer Antibiotics." Ph.D Biochemistry, University of Colorado at Boulder (Marvin Caruthers). Brummel's thesis work has been published, publication #24. Currently Research Scientist at Agilent Tech in the Bay Area.

Karl Kirschner - Honors Senior Thesis, "Quantum Mechanical Investigation of Adenosine 3',5'-Cyclic Monophosphate." Ph.D. Physical Chemistry, University of Georgia (Phillip Bowen). Kirschner has published four papers based on his undergraduate research, publications #19, #23, #27, & #28. Currently Visiting Assistant Professor of Chemistry at Hamilton College.

Tricia Lively - Honors Senior Thesis, "Design of an Expression System for *Chlorella* XZ-6E **CviRI** Methylase, Protein Purification of **CviRI** Methylase, *Providencia stuartii* 164 **PstI** Endonuclease and Methylase Proteins, and Crystallization of **PstI** Endonuclease with a 20 Basepair Sequence of DNA." Ph.D. completed at the University of Colorado at Boulder (Jim Goodrich). Lively has published two papers from her undergraduate research, publications #22 & #29. Currently postdoc with Erwin Gelfand at the National Jewish Medical and Research Center in Denver.

1993

Roger Sommer (co-advised with M.L. Thompson) - Honors Senior Thesis, recipient of the *Phi Beta Kappa Senior Thesis Award* "Computational and Experimental Approaches to the Synthesis of Polycyclic Phosphazanes." Ph.D. Inorganic Chemistry, University of Colorado at Boulder (Gordon Yee). Postodoctoral associate at the University of Delaware, (Arnie Rheingold); taught at Middlebury College; currently Assistant Professor of Chemistry at Loyola University.

1992

Ashish Desai - Honors Senior Thesis "Evolution of Lysozyme and Alpha-lactalbumin." MS Chem., University of Illinois at Chicago (Albert Benight). MBA, Texas A&M.

Graham Haas - Senior Project "Hydrogen Bonding of Formic Acid: A FT-IR Study." MS Chem., Georgia Tech (Ken Busch). Employed in industry.

Sarah Gallagher – Summer Research in Thermodynamics. Ph.D. Physical Chemistry, University of Colorado at Boulder (David Jonas).

1991

Kris Henningfeld - Honors Senior Thesis "Purification of Glyceraldehyde-3-Phosphate Dehydrogenase from Chicken Muscle." Ph.D. Chem., University of Virginia (Sidney Hecht). Currently postdoctoral associate in the Department of Oncology/Hematology at the University of Ulm, Germany.

Tracey Holmes - Honors Senior Thesis "Purification and Characterization of Pigeon Intestine Alkaline Phosphatase." Currently research and development technician at Abbott Laboratories.

Marcus Jurema - Honors Senior Thesis, co-recipient of the *Phi Beta Kappa Senior Thesis Award* "An Investigation of the Applicability of the PM3 Semi-empirical Quantum Mechanical Method to Hydrogen Bonded Systems." M.D. Stanford University; further training at Cornell Medical School and Johns Hopkins Medical School. Jurema's thesis work was published in three publications, publications #18 and #19, and #22. Currently on faculty at Brown Medical School as a Clinical Assistant Professor of Reproductive Medicine and Infertility.

1990

Georgia Kalkanis - Honors Senior Thesis "Quantum Mechanical Calculations of the Reactions of CH_2OH with O_2, NO and NO_2." D.D.S., University of Detroit. In private practice in Detroit. Kalkanis's thesis work was published in 1991, publication #16.

INVITED LECTURES

23.) "Computational Design of a Small Peptide that Inhibits Breast Cancer." Memorial Sloan-Kettering Cancer Center, September 12, 2006. Audience 40.

22.) "Computational design of a small peptide that inhibits breast cancer: An overview of computational chemistry research at Hamilton College." 38th ACS Middle Atlantic Regional Meeting, June 5, 2006. Audience 25.

21.) "Computational Design of a Small Peptide that Inhibits Breast Cancer." Florida State University, March 23, 2006. Audience 30.

20.) "Computational Design of a Small Peptide that Inhibits Breast Cancer" at the 46th International Symposium on Atomic, Molecular, Biophysical, and Condensed Matter Theory, March 6, 2006, St. Simons Island, GA. Audience 125.

19.) "Water Clusters in the Atmosphere: An Overview of Computational Chemistry Research at Hamilton College." Ithaca College, November 15, 2005. Audience 25.

18.) "Water Clusters in the Atmosphere: An Overview of Computational Chemistry Research at Hamilton College." Syracuse University, September 20, 2005. Audience 30.

17.) "Water Clusters in the Atmosphere: An Overview of Computational Chemistry Research at Hamilton College." University of Minnesota, Minneapolis, December 14, 2004. Audience 100, plus seminar was webcast to 35 other colleges in the RSEC program.

16.) "Hybrid QM/MM Models" at the 44th International Symposium on Atomic, Molecular, Biophysical, and Condensed Matter Theory, March 2004, St. Augustine, Florida. Audience 30.

15.) "Formation of MERCURY to Enhance Undergraduate Computational Chemistry: Accurate pK_a Calculations in Aqueous Solution, Progress & Challenges" at the 35th ACS Central Regional Meeting, Pittsburgh, PA, October 21, 2003. Audience 35.

14.) "An Overview of Chemical Physics Research at Hamilton College: Accurate pK_a Calculations in Aqueous Solution, Progress and Challenges." University of Maryland, College Park, April 16, 2003. Audience 40.

13.) "Career Options in the Field of Chemistry." Clinton Central High School, two Honors Chemistry Classes, December 4, 2002. Audience 24.

12.) "Basic Research and Drug Design: Computational Chemistry at Hamilton College." Wright State University, October 18, 2002. Audience 32.

11.) "Formation of MERCURY and Acquisition of High-Performance Computers to Support Research in Undergraduate Computational Chemistry across New York and New England" at the Executive Summit, *Delivering Technology Leadership for Life Sciences: Research Advances for Drug Discovery and Bioterrorism.* A Life Sciences Executive Summit co-sponsored by SGI, the Delaware Biotechnology Institute (DBI), and the University of Delaware. DBI, October 3, 2002. Audience 100.

10.) "Basic Research and Drug Design: Computational Biochemistry at Hamilton College." Union College, February 21, 2002. Audience 40.

9.) "How to Succeed in College" Marcellus Academic Club Honors Recognition Dinner, Marcellus, N.Y., March 25, 2001. Audience 100.

8.) "One Semester, Accelerated General Chemistry at Hamilton College." St. Lawrence University, December 7, 2000. Audience 6.

7.) "An Overview of Computational Chemistry at Hamilton College: Drug Design, the Nature of Water Dissociation, and Accurate pK_a Calculations." St. Lawrence University, December 7, 2000. Audience 12.

6.) "An Overview of Computational Chemistry at Hamilton College: Drug Design, the Nature of Water Dissociation, and Accurate pK_a Calculations." Vassar College, November 29, 2000. Audience 32.

5.) "Chemical Physics in the Chemistry Department." Hamilton College, Department of Physics, November 6, 2000. Audience 25.

4.) "Computational Chemistry at Hamilton College: Drug Design, the Nature of Water Dissociation, and Accurate pK_a Calculations." Connecticut College, October 24, 2000. Audience 32.

3.) "Computational Chemistry at Hamilton College: Drug Design, the Nature of Water Dissociation, and pK_a Calculations." Holy Cross College, September 29, 2000. Audience 34.

2.) "An Overview of Computational Chemistry at Hamilton College. Can quantum chemistry be used to calculate accurate pK_a values?" Hobart & William Smith Colleges, June 2, 2000. Audience 65.

1.) "Formation of Base Pairs Containing Adenine, Thymine, and difluorotoluene: The Role of Solvent," at the ISQBP President's Meeting on Molecular Structure and Dynamics in Biology, September 1998, Elba, Italy. Audience 250.

PROFESSIONAL DEVELOPMENT & RECENT ACTIVITIES

July 2006 Organized National Conference, "5th MERCURY Conference in Computational

	Chemistry."
March 2006	Met with Ouathek Ouerfelli and Yueming Li at Memorial Sloan-Kettering Cancer Center to begin collaborations in cancer biochemistry.
Novemb er 2005	Met with Gavril Pasternak at Memorial Sloan-Kettering Cancer Center to discuss research in cancer biochemistry.
October 2005	Met with Veronica Vaida at the University of Colorado and colleagues at NCAR to develop collaboration in Atmospheric Chemistry
July 2005	Met with Tom Andersen (Albany Medical College) to suggest a molecule based on AFP as a potential drug for breast cancer
July 2005	Organized National Conference, "4th MERCURY Conference in Computational Chemistry."
July 2004	Organized National Conference, "3rd MERCURY Conference in Computational Chemistry."
July 2003	Organized National Conference, "2nd MERCURY Conference in Computational Chemistry."
July 2002	Organized National Conference, "1st MERCURY Conference in Computational Chemistry."
April 2002	Advances in Computational Drug Discovery, workshop by Accelrys, Boston.
March 2002	Met with collaborator Tom Andersen to discuss progress on AFP project.
September 2001	Visited SGI, in Mountain View California, for discussions of high performance computers.
April 2001	Met with Tom Andersen (Albany Medical College) to develop a collaboration around AFP, a protein that protects against breast cancer

PROFESSIONAL SOCIETIES

American Association for the Advancement of Science
American Chemical Society - Physical Chemistry Division, Theoretical Chemistry Division, Computers in Chemistry Division
Council on Undergraduate Research
International Society of Quantum Biology and Pharmacology
Sigma Xi

CURRENT FIELDS OF INTEREST

Computational Chemistry; Structural Biochemistry; Atmospheric Chemistry; Undergraduate Research; Science Education

MAJOR COMMUNITY SERVICE

Hamilton College

Member of the President's advisory committee, 2005-2006
Clinton High School Committees: Safe Routes to School; Class Rank Review, 2006
Chair, Chemistry Department, 1999-2006
Chair, Chemical Physics Program, 2000-06
Acting Chair, Biochemistry & Molecular Biology Program, 2004-06
Committee on Athletics (elected by faculty) 2002
Committee on Academic Appointments (elected by faculty to three year term) 2000-02
Member of the Architect Selection committee for the new Science Center building, 2000
Member of the Biochemistry/Molecular Biology Program, 1998-present
Member of the President's advisory committee, 1999-2001

Lake Forest College

Chair, Academic Resources and Review Committee 1997-98
Academic Resources and Review Committee (elected by faculty, 1 year & 3 year terms) 1996-97, 1997-98
College Council (President's Advisory Committee, elected by faculty) 1991-1994, 1994-1995
Budget and Audit Committee of the Board of Trustees 1994-95
Pew Roundtable Participant 1994
Faculty Representative on the Board of Trustees 1993-1994
Search Committee for Provost and Dean of the Faculty 1993
President, LFC Chapter of Sigma Xi 1993-1994
Program Chair, LFC Chapter of Sigma Xi 1992-1993
Ad hoc Committee on Course and Teaching Credit - major author of variable credit proposal 1991
Search Committee for Director of Information Services and Technology 1990-1991

PEER REVIEWED PUBLICATIONS (Undergraduate Co-authors denoted by *)

57. "Enediyne Bergman Cyclization of the Chemically Active Region of Esperamicin A_1" Frank C. Pickard IV,* Chantelle Rein,* Steven Feldgus, Karl N. Kirschner and George C. Shields, in preparation for *Journal of Physical Chemistry B* .

56. "Computational Design and Experimental Discovery of an Anti-estrogenic Peptide Derived from Alpha-Fetoprotein" Karl N. Kirschner, Katrina W. Lexa,* Amanda M. Salisburg,* Katherine A. Alser,* Leroy Joseph, Thomas T. Andersen, James A. Bennett, Herbert I. Jacobsen, and George C. Shields, in preparation for **Nature**.

55. "The Nature of the Dissociation of H_2O in Water" Sarah T. Tschampel,* Frank C. Pickard IV,* Jennifer R. Derby,* and George C. Shields, in preparation.

54. "In Search of $CS_2(H_2O)_{n=1-4}$ Clusters" Gregory M. Hartt,* Timothy M. Evans,* Karl N. Kirschner and George C. Shields, **J. Chem. Phys.** submitted 9/6/06, in revision.

53. "Do Hydroxyl Radical-Water Clusters, $OH(H_2O)_n$, n=1-5, Exist in the Atmosphere?" Marco A. Allodi,* Meghan E. Dunn,* Jovan Livada,* Karl N. Kirschner and George C. Shields, **J. Phys. Chem. A** accepted for publication 9/28/2006.

52. "Comparison of CCSD(T), W1, and other Model Chemistry Predictions for Gas Phase Deprotonation Reactions" Frank C. Pickard IV,* Daniel R. Griffith,* Skylar J. Ferrara,* Matthew D. Liptak,* Karl N. Kirschner and George C. Shields, **Int. J. Quantum Chem. 106** (2006) 3122-3128.

51. "Exploration of the Potential Energy Surfaces, Prediction of Atmospheric Concentrations, and Vibrational Spectra of the $HO_2\bullet\bullet\bullet(H_2O)_n$ (n=1-2) Hydrogen Bonded Complexes" Kristin S. Alongi,* Theodore S. Dibble, George C. Shields and Karl N. Kirschner, **J. Phys. Chem. A 110** (2006) 3686-3691.

50. "Ortho-Effect in the Bergman Cyclization: Electronic and Steric Effects in Hydrogen Abstraction by 1-Substituted Naphthalene 5,8-Diradicals" Frank C. Pickard IV,* Rebecca L. Shepherd,* Amber E. Gillis,* Meghan E. Dunn,* Steven Feldgus, Karl N. Kirschner, George C. Shields, Mariappan Manoharan, and Igor V. Alabugin, **J. Phys. Chem. A 110** (2006) 2517-2526.

49. "Prediction of Accurate Anharmonic Experimental Vibrational Frequencies for Water Clusters, $(H_2O)_n$, n=2-5" Meghan E. Dunn,* Timothy M. Evans,* Karl N. Kirschner and George C. Shields,

J. Phys. Chem. A 110 (2006) 303-309.

48. "Comparison of Density Functional Theory Predictions of Gas Phase Deprotonation" Matthew D. Liptak* and George C. Shields, **Int. J. Quantum Chem. 105** (2005) 580-587.

47. "Global Search for Minimum Energy $(H_2O)_n$ Clusters, n=3-5" Mary Beth Day,* Karl N. Kirschner and George C. Shields, **J. Phys. Chem. A 109** (2005) 6773-6778.

46. "Comparison of Model Chemistry and Density Functional Theory Thermochemical Predictions with Experiment for Formation of Ionic Clusters of the Ammonium Cation Complexed with Water and Ammonia; Atmospheric Implications." Frank C. Pickard IV,* Meghan E. Dunn* and George C. Shields, **J. Phys. Chem. A 109** (2005) 4905-4910.

45. "First-principle studies of intermolecular and intramolecular catalysis of protonated cocaine" Chang-Guo Zhan, Shi-Xian Deng, Jaime G. Skiba,* Beth A. Hayes,* Sarah M. Tschampel,* George C. Shields and Donald W. Landry, **J. Comput. Chem. 26** (2005) 980-986.

44. "Pople's Gaussian-3 model chemistry applied to an investigation of $(H_2O)_8$ water clusters" Mary Beth Day,* Karl N. Kirschner and George C. Shields, **Int. J. Quantum Chem. 102** (2005) 565-572.

43. "Comparison of CBSQB3, CBSAPNO, G2, and G3 thermochemical predictions with experiment for formation of ionic clusters of hydronium and hydroxide ions complexed with water" Frank C. Pickard IV,* Emma K. Pokon,* Matthew D. Liptak* and George C. Shields, **J. Chem. Phys. 122** (2005) 024302.

42. "The ability of the Gaussian-2, Gaussian-3, Complete Basis Set-QB3, and Complete Basis Set-APNO model chemistries to model the geometries of small water clusters" Meghan E. Dunn,* Emma K. Pokon* and George C. Shields, **Int. J. Quantum Chem. 100** (2004) 1065-1070.

41. "Accurate Experimental Values for the Free Energies of Hydration of H^+, OH^-, and H_3O^+" Matthew W. Palascak* and George C. Shields, **J. Phys. Chem. A 108** (2004) 3692-3694.

40. "Thermodynamics of Forming Water Clusters at Various Temperatures and Pressures by Gaussian-2, Gaussian-3, Complete Basis Set-QB3, and Complete Basis Set-APNO Model Chemistries; Implications for Atmospheric Chemistry" Meghan E. Dunn,* Emma K. Pokon,* and George C. Shields, **J. Am. Chem. Soc. 126** (2004) 2647-2653.

39. "Absolute pK_a Determinations for Substituted Phenols" Matthew D. Liptak,* Kevin C. Gross, Paul G. Seybold, Steven Feldgus, and George C. Shields, **J. Am. Chem. Soc. 124** (2002) 6421-6427.

38. "Comparison of CBS-QB3, CBS-APNO, and G3 predictions of gas phase deprotonation data" Emma K. Pokon,* Matthew D. Liptak,* Steven Feldgus, and George C. Shields, **J. Phys. Chem. A 105** (2001) 10483-10487.

37. "An ONIOM Study of the Bergman Reaction: A Computationally Efficient and Accurate Method for Modeling the Enediyne Anticancer Antibiotics" Steven Feldgus and George C. Shields, **Chem. Phys. Lett. 347** (2001) 505-511.

36. "Experimentation with different thermodynamic cycles used for pK_a calculations on carboxylic acids using Complete Basis Set and Gaussian-n Models combined with CPCM Continuum Solvation Methods" Matthew D. Liptak* and George C. Shields, **Int. J. Quantum Chem. 85**

(2001) 727-741.

35. "Further Quantum Mechanical Evidence that Difluorotoluene does not Hydrogen Bond" Edward C. Sherer, Sam J. Bono,* and George C. Shields, **J. Phys Chem. B 105** (2001) 8445-8451.

34. "Accurate pK_a Calculations for Carboxylic Acids Using Complete Basis Set and Gaussian-n Models Combined with CPCM Continuum Solvation Methods" Matthew D. Liptak* and George C. Shields, **J. Am. Chem. Soc. 123** (2001) 7314-7319.

33. "Accurate relative pK_a calculations for carboxylic acids using complete basis set and Gaussian-n models combined with continuum solvation methods" Ann Marie Toth,* Matthew D. Liptak,* Danielle L. Phillips,* and George C. Shields, **J. Chem. Phys. 114** (2001) 4595-4606.

32. "Molecular Dynamics Simulation of a PNA·DNA·PNA Triple Helix in Aqueous Solution" George C. Shields, Charles A. Laughton, and Modesto Orozco, **J. Am. Chem. Soc. 120** (1998) 5895-5904.

31. "Comparison of Experimental and Theoretical Structures of a Transition State Analog used for the Induction of Catalytic Antibodies that Destroy Cocaine" Edward C. Sherer,* Ginger Yang, Gordon M. Turner,* George C. Shields and Donald W. Landry, **J. Phys. Chem. A** 101 (1997) 8526-8529.

30. "Molecular Dynamics Simulations of the d(T·A·T) Triple Helix" George. C. Shields, Charles A. Laughton, and Modesto Orozco, **J. Am. Chem. Soc. 119** (1997) 7463-69.

29. "A Semiempirical Transition State Structure for the First Step in the Alkaline Hydrolysis of Cocaine. Comparison between the Transition State Structure, the Phosphonate Monoester Transition State Analog, and a Newly Designed Thiophosphonate Transition State Analog." Edward C. Sherer*, Gordon M. Turner,* Tricia N. Lively,* Donald W. Landry and George C. Shields, **J. Mol. Model. 2** (1996) 62-69.

28. "Use of the Supermolecule Approach to Model the *syn* and *anti* Conformations of Solvated Cyclic 3',5'-Adenosine Monophosphate" Karl N. Kirschner,* Edward C. Sherer,* and George C. Shields, **J. Phys. Chem. 100** (1996) 3293-3298.

27. "Quantum Mechanical Investigation of Cyclic 3',5'- Adenosine Monophosphate, the Second Hormonal Messenger" Karl N. Kirschner* and George C. Shields, **J. Mol. Structure (THEOCHEM) 362** (1996) 297-304.

26. "A Computationally Efficient Procedure for Modeling the First Step in the Alkaline Hydrolysis of Esters" Gordon M. Turner,* Edward C. Sherer,* and George C. Shields, **Int. J. Quantum Chem. 56** (1995) 103-112.

25. "Investigation of the Potential Energy Surface for the First Step in the Alkaline Hydrolysis of Methyl Acetate" Edward C. Sherer*, Gordon M. Turner,* and George C. Shields, **Int. J. Quantum Chem. 56** (1995) 83-93.

24. "Semiempirical Study of the Bergman Reaction: Towards a Computationally Efficient and Accurate Method for Modeling the Enediyne Anticancer Antibiotics" Heather A. Brummel* and George C. Shields, **Int. J. Quantum Chem. 56** (1995) 51-59.

23. "Quantum-Mechanical Investigation of Large Water Clusters" Karl N. Kirschner* and George C. Shields, **Int. J. Quantum. Chem. 52** (1994) 349-360.

22. "Hydrogen Bonding of Nucleotide Base Pairs: Application of the PM3 Method" Tricia N. Lively,* Marcus W. Jurema,* and George C. Shields, **Int. J. Quantum Chem. 52** (1994) 95-107.

21. "The Physical Chemistry Sequence at Liberal Arts Colleges: The Lake Forest College Approach" George C. Shields, **J. Chem. Educ. 71** (1994) 951-953.

20. "Using the Franck-Hertz Experiment to Illustrate Quantization: Energy States of the Neon Atom by Electron Impact" Michael M. Kash and George C. Shields, **J. Chem. Educ. 71** (1994) 466-468.

19. "Modeling of Magic Water Clusters $(H_2O)_{20}$ and $(H_2O)_{21}H^+$ with the PM3 Quantum-Mechanical Method" Marcus W. Jurema,* Karl N. Kirschner,* and George C. Shields, **J. Comput. Chem. 14** (1993) 1326-1332.

18. "Ability of the PM3 Quantum Mechanical Method to Model *Inter*molecular Hydrogen Bonding between Neutral Molecules" Marcus W. Jurema* and George C. Shields, **J. Comput. Chem. 14** (1993) 89-104.

17. "Experiment in Quantization: Atomic Line Spectra" George C. Shields and Michael M. Kash, **J. Chem. Educ. 69** (1992) 329-331.

16. "AM1 and PM3 Calculations of the Potential Energy Surfaces for CH_2OH Reactions with NO and NO_2" Georgia H. Kalkanis* and George C. Shields, **J. Phys. Chem. 95** (1991) 5085-5089.

15. "Crystal Structure of a CAP-DNA Complex Shows that the DNA is bent 90°" Steve C. Schultz, George C. Shields and Thomas A. Steitz, **Science 253** (1991) 1001-1007.

14. "Crystallization of *Escherichia coli* Catabolite Gene Activator Protein with its DNA Binding Site: The use of Modular DNA" Steve C. Schultz, George C. Shields and Thomas A. Steitz, **J. Mol. Biol. 213** (1990) 159-166.

13. "Structural Studies of Three DNA Binding Proteins: Catabolite Gene Activator Protein, Resolvase, and the Klenow Fragment of DNA Polymerase I" T.A. Steitz, L. Beese, B. Engelman,* P. Freemont, J. Friedman, M. Sanderson, S. Schultz, G. Shields and J. Warwicker, **NATO ASI Ser., Ser. A 137** (1987) 185-189.

12. "Competition between Single and Double Electron Transfer in Collisions of Doubly Charged Molecular Pyrrole Ions with Neutral Pyrrole Molecules" J.B. Sedgwick, B.P. Paulson, G.C. Shields and T.F. Moran, **Int. J. Mass Spectrom. Ion Proc. 79** (1987) 127-140.

11. "Doubly Charged Ion Mass Spectra of Alkyl Substituted Furans and Pyrroles" P.R. Nelson, C.A. Fung Kee Fung, J.B. Sedgwick, G.C. Shields, L.E. Abbey and T.F. Moran, **Org. Mass Spectrom. 22** (1987) 389-399.

10. "*Charge Transfer Reactions of Organic Ions Containing Oxygen:* Correlation Between Reaction Energetics and Cross Sections" G.C. Shields and T.F. Moran, **Org. Mass Spectrom. 21** (1987) 64-69.

9. "*Doubly Charged Ethane Ions:* Solution to the Dilemma of Stability Predicted by Theory and Instability Observed in Experiment" G.C. Shields and T.F. Moran, **Org. Mass Spectrom. 21** (1986) 479-483.

8. "Polarizabilities of Organic Ions" G.W. Burdick,* G.C. Shields and T.F. Moran , **Org. Mass Spectrom. 21** (1986) 449-450.

7. "Doubly Charged Gas Phase Cations" G.C. Shields and T.F. Moran, **Theoretica Chimica Acta 69** (1986) 147-159.

6. "Sensitivity of Charge Transfer Reactions to Hydrocarbon Ion Structures" G.C. Shields, L. Wennberg,* J.B. Wilcox and T.F. Moran, **Org. Mass Spectrom. 21** (1986) 137-149.

5. "Evidence for Long-lived Excited States of $C_nH_2^{2+}$ Carbodications" J.R. Appling, G.C. Shields and T.F. Moran, **Org. Mass Spectrom. 21** (1986) 69-75.

4. "Molecular Charge Transfer Cross Sections and Their Correlation with Reactant Ion Structures" G.C. Shields and T.F. Moran, **J. Phys. Chem. 89** (1985) 4027-4031.

3. "Structures, Energetics and Fragmentation Pathways of $C_nH_2^{2+}$ Carbodications " G.W. Burdick,* G.C. Shields, J.R. Appling and T.F. Moran, **Int. J. Mass Spectrom. Ion Proc. 64** (1985) 315-333.

2. "Double Electron Transfer Reactions of CO_2^{2+} Ions" G.C. Shields and T.F. Moran, **Chem. Phys. Lett. 101** (1983) 287-290.

1. "Single- and Double-Electron Transfer Reactions of Ground and Metastable State Ar^{2+} Ions" G.C. Shields and T.F. Moran, **J. Phys. B: At. Mol. Phys. 16** (1983) 3591-3601.

Since 1983 I have co-authored 53 papers. I have published 32 papers with 27 undergraduates working in my research group since 1991; 35 papers with 29 undergraduates total.

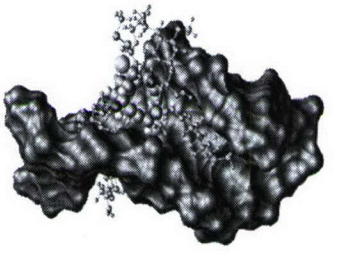
OTHER PUBLICATIONS

1. "The Benefits of Forming a Consortium for an NSF-MRI Proposal" George C. Shields, CUR Quarterly, 23 (2002) 80-81.

Enediyne Bergman Cyclization of the Chemically Active Region of Esperamicin A_1

Frank C. Pickard IV, Chantelle Rein, Steven Feldgus,
Karl N. Kirschner and George C. Shields*
Department of Chemistry, Hamilton College, Clinton, NY 13323

*gshields@hamilton.edu



Abstract

Esperamicin A_1 is a member of a larger family of naturally occurring products that share a novel molecular architecture and potent biological activity. The purpose of this study is to derive and evaluate an effective model compound and methodology for the computational study of the Bergman cyclization in this naturally occurring product. The model compound presented in this paper, MTC2, achieves a significant reduction in size by omitting the chemically inert oligosaccharide arms that are crucial for esperamicin A_1's ability to recognize and dock to DNA. We present an effective computational methodology for the study of Bergman cyclization in small and moderate sized enediyne molecules.

1. Introduction

Esperamicin A_1 (Figure 1) is a member of a larger family of naturally occurring products that share a novel molecular architecture and potent biological activity. Other members of this family include: dynemicin A, neocarzinostatin and calicheamicin $_{\gamma 1}$. The

esperamicins were discovered in 1985,[1] isolated from a fermentation broth of *actinomadura verrucosospora*. Esperamicin was named for the place where the bacteria responsible for the production of this compound were collected, Pto Esperanza, in Argentina. In 1987 the structures of the esperamicins were reported.[2,3] All four of these enediyne compounds have drawn significant interest as potential sources of anticancer drugs because of their ability to cleave DNA, presumably through a Bergman cyclization.[4] These compounds are among the world's most powerful anti-tumor agents, as they exhibit high activity against a variety of murine tumor models at injected doses in the 0.1 $\mu g \cdot kg^{-1}$ range.[5,6] However, the enediyne natural products are not selective in their activity, and will cleave DNA in both healthy and cancerous cells. This indiscriminant behavior has spurred a flurry of both experimental[7-11]and computational[12-18] research focused upon rational drug design featuring the enediyne moiety. However, the exact mechanism responsible for esperamicin A_1's cytotoxicity is not thoroughly understood. This paper is an initial step towards understanding how this natural product may cleave DNA at a molecular level.

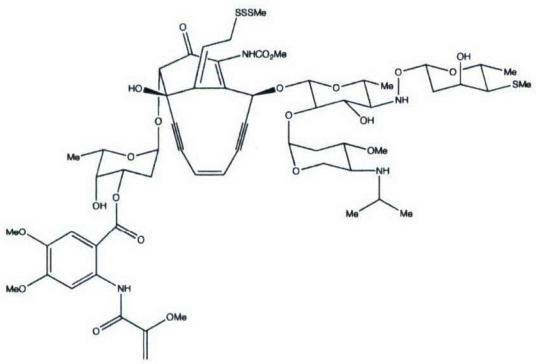


Figure 1. The structure of the untriggered esperamicin A_1 molecule.

In the most commonly accepted mechanism for esperamicin's efficacy, a bioreductuve cleavage of the trisulfide tail is carried out by glutathione generating a thiol.[19] The thiol then undergoes an intramolecular addition to the 10-membered ring containing the enediyne moiety, reducing strain and increasing flexibility in the enediyne ring,[20-22] which significantly lowers the activation barrier for Bergman cyclization (Figure 2). This cyclization reaction is an electronic rearrangement reaction where the 1,5-diyne-3-ene moiety rearranges to form a 1,4-didehydrobenzene diradical. Abstraction of hydrogen atoms from the DNA backbone by the diradical leads to single-stranded DNA cleavage and eventual apoptosis. Recently, Alabugin *et al.*[23] proposed an alternate mechanism for radical hydrogen abstraction. They argue for the feasibility of an intermediate intramolecular hydrogen abstraction step that would occur after Bergman cyclization and before the hydrogen abstraction from DNA. This intramolecular abstraction would essentially 'quench' the *p*-benzyne intermediate and make deactivation pathways such as the retro Bergman cyclization inaccessible, increasing the lifespan of the DNA damaging species.

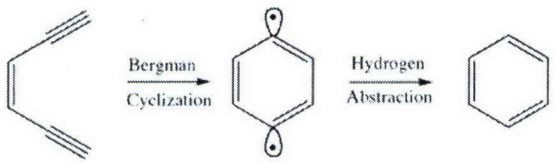


Figure 2. Schematic of the Bergman Cyclization and hydrogen abstraction.

The purpose of this study is to derive and evaluate an effective model compound and methodology for the computational study of the Bergman cyclization in naturally occurring products, specifically esperamicin A_1. Because of the extensive size of the

esperamicin A_1 molecule, and the rigorous theoretical treatment that its multi-reference electronic state demands, we have chosen to work with a series of smaller models that contain esperamicin A_1's chemically active core. The model compound presented in this paper, MTC2 (Figure 3), is a variation on the smaller MTC chemical model previously used by our group to assist in the study of the Bergman cyclization.[24] The MTC2 model achieves a significant reduction in size by omitting the chemically inert oligosaccharide arms that are important for esperamicin A_1's docking to DNA. Also, since this study is not concerned with the triggering mechanism, the nitrogen tail, which is adjacent to the five-membered sulfur heterocycle, has also been pruned to further reduce the size of the system. These omissions reduce the system size from one hundred sixty atoms to forty atoms, yet the MTC2 model incorporates esperamicin A_1's multi-cyclic ring structure and neighboring substituent groups into the calculations.

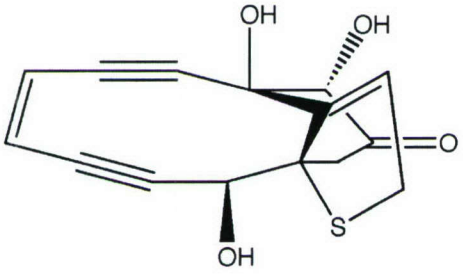


Figure 3. The MTC2 model system used in this study. Hydroxyl groups have replaced the carbohydrates and a hydrogen-capping atom has replaced the nitrogen tail.

2. Methods

The Bergman cyclization is a difficult reaction to model because of the multi-configurational nature of the diradical species. Reliable wavefunction descriptions of the diradical state require extensive incorporation of dynamic electron correlation effects.

Both the UHF-CC and RB-CC levels of theory have been shown to provide accurate diradical state descriptions in *p*-benzyne,[25] but these methods are prohibitively expensive for all but the smallest enediyne model systems. Semiempirical theory has shown some promise, yet it is difficult to benchmark these methods without experimental results on the natural products themselves.[26] Density functional theory (DFT) has emerged as a promising high-accuracy, low-cost alternative to high order dynamic correlation methods. In the past five years broken-symmetry (BS-UDFT) approaches have emerged as an acceptable alternative to dynamic correlation methods. While DFT has been widely lauded for its ability to reproduce experimental enthalpies for a variety of small enediynes, it is important to note that DFT's ability to accurately model spin effects, magnetic properties and other properties that depend upon spin is still in question.[27]

Because of the open shell nature of both the diradical and transition state species, all calculations on these structures were performed as BS-UDFT calculations. Both the "Guess=Mix" and "Stable=Opt" keywords were used to generate open shell singlet wavefunctions. We obtained two sets of optimized geometries for this system, using the B3LYP and the PBE1PBE (PBE0) density functionals with the 6-31G(d,p) basis set. Frequency calculations were carried out upon all optimized structures to ensure their true minimization and to provide thermochemical corrections to computed electronic energies. All thermochemistry data is given at a standard state of one atmosphere. The pre-cyclized and diradical structures represent local minima on the potential energy surface, and have no imaginary frequencies. The optimized transition state structures have one imaginary frequency along the mode corresponding to the bond formation in the Bergman Cyclization reaction. A series of single point calculations were carried out on

both sets of optimized structures in the gas phase using the following basis sets: 6-311G(2d,p), cc-pVDZ, aug-cc-pVDZ, cc-pVTZ and aug-cc-pVTZ.[28-31] Tight SCF convergence criteria were requested for all calculations. Solvation was modeled implicitly using a polarizable continuum model (IEFPCM)[32-34] along with a molecular cavity optimized for use at the HF/6-31G* level of theory (UAHF model). All computations were performed using revisions B.02 and C.02 of Gaussian 03.[35]

3. Results & Discussion

The most challenging aspect of modeling this system is accurately and reliably describing the diradical product molecule. This challenge is two-fold: finding an acceptable single determinant method for describing the multi-reference diradical state, and converging to the correct open shell state. Because of the quasi-degeneracy of the two outermost occupied orbitals[25] it is easy to converge to an incorrect closed shell state. Thus, we need a metric to determine if our converged electronic state is truly the most stable one.

3.1 Wavefunctions

As stated in the methods section, BS-UDFT was used to produce open shell singlet states for the diradical species; unfortunately this method was unreliable. While usually effective, there were instances where the broken symmetry wavefunction still collapsed to a closed shell state, resulting in poor energies (errors exceeding 20 kcal•mol^{-1}). We ascertained whether a state properly converged to an open shell solution by inspection of the $\langle S^2 \rangle$ values. For transition states, we found that the most stable open shell states had a spin excess of approximately 0.1 before spin annihilation. We also

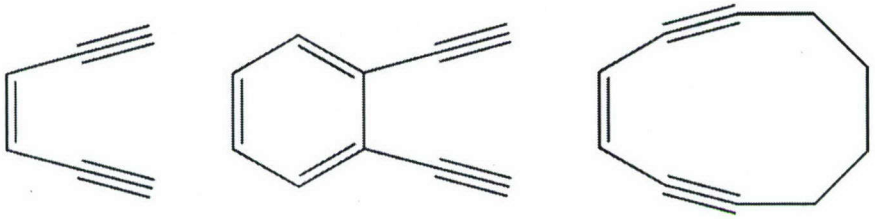
found that the most stable open shell diradical wavefunctions had a spin excess of approximately 1 prior to spin annihilation. This result is qualitatively correct because the erroneous closed shell wavefunctions will have no triplet contamination and therefore spin symmetry breaking is not taking place. The empirical accuracy of $\langle S^2 \rangle$ values predicted by DFT has been a topic of some debate.[36] We are not interested in the $\langle S^2 \rangle$ values in and of themselves; we are only interested in them as a metric for determining whether we have in fact converged to an open shell solution.

In the cases where BS-UDFT seemed to fail to produce open shell states, we tested the stability of the solutions via the "Stable=Opt" keyword. This routine tests the stability of a wavefunction using Hermitian stability matrices, and if the wavefunction is not stable, re-optimizes it by permuting the orbital space of the original solution as dictated by the eigenvectors of the stability matricies.[37] While this method is very reliable, it is extremely expensive, and thus its utility is limited to small systems. As long as the $\langle S^2 \rangle$ values are rigorously checked, BS-UDFT is a fine method for generating proper states. In the rare cases where BS-UDFT fails, and stability calculations are prohibitively expensive, we found that a combination of symmetry breaking and valence orbital shifting to be useful. While this method is also unreliable, the calculations are quick enough that a 'guess and check' approach is feasible, using the $\langle S^2 \rangle$ values as the diagnostic tool.

3.2 Benchmarking Calculations

In previous work[24] we made benchmark calculations on three small enediyne molecules: *cis*-hex-3-ene-1,5-diyne (**1**), *o*-diethynylbenzene (**2**) and cyclodec-3-ene-1,5-

diyne (**3**), shown in Figure 4. These three enediynes are particularly important because experimental data is available for both enthalpies of reaction and enthalpies of activation.



1 **2** **3**

Figure 4. Three small enediyne molecules for which accurate experimental data is available.

Calculated enthalpies from our previous work,[24] from the work of Schreiner *et al.*,[38] and experimental values[39-41] are presented in Table 1. A few interesting trends emerge. First, enthalpies of reaction, calculated using the B3LYP functional, improve with relation to experiment as the basis set is expanded from double-zeta to triple-zeta. Second, enthalpies of reaction are consistently underestimated for double-zeta calculations for both **1** and **2**. Conversely, enthalpies calculated at the triple-zeta level are slightly overestimated. Third, for the two pure DFT functionals, BLYP and G96LYP, the opposite trend emerges. Enthalpies calculated with the smaller double-zeta basis sets follow experimental results more closely, and triple-zeta calculations result in much larger errors. This result implies that pure density functionals may be inferior to hybrid density functionals for modeling *p*-benzyne diradicals and their analogues accurately.[42] Fourth, enthalpies of activation show reverse trends from the enthalpies of reaction. Double-zeta B3LYP calculations appear to be highly accurate predictors of $\Delta H^{\ddagger}$, while the triple-zeta calculations overestimate the barrier by 4-6 kcal•mol^{-1}. Both pure

functionals consistently underestimate the barrier height, though expansion of the basis set from double- to triple-zeta reduces the errors from ~4 to ~1 kcal•mol^{-1}.

From our reevaluation of our previous benchmarking data and the calculations performed by Schreiner *et al.*, we decided to employ the B3LYP hybrid functional for studying the MTC2 model system. Because the errors relative to experimental enthalpies of reaction increase when the basis set is expanded when using the pure functionals, the good agreement between the double-zeta results and experiment is likely due to a fortuitous cancellation of errors. Also, the use of a sum formula was shown to dramatically improve results calculated with B3LYP, but not for results calculated with pure functionals.[42,43] This formula estimates the energy of a pure singlet state by assuming that all spin contamination is introduced by the first triplet state. The energy contributions from this triplet state are then estimated and removed.

Table 1. Enthalpies of activation and reaction for three benchmark molecules, deviations from experiment are italicized. All values are in kcal•mol^{-1}.

		1				**2**				**3**		
		ΔH		$\Delta H^{\ddagger}$ (470 K)		ΔH		$\Delta H^{\ddagger}$ (486 K)		ΔH	$\Delta H^{\ddagger}$ (298 K)	
B3LYP	6-31G(d)[b]	3.94	*-4.6*	29.31	*1.1*							
	6-31G(d,p)[a]	3.0	*-5.5*	29.4	*1.2*	12.0	*-5.8*	29.5	*4.3*	3.3	26.0	*2.8*
	6-31++G(d,p)[b]	5.90	*-2.6*	29.39	*1.2*							
	cc-pVDZ[b]	4.33	*-4.2*	28.95	*0.8*							
	6-311G(d,p)[b]	10.34	*1.8*	32.42	*4.2*							
	cc-pVTZ[a]	10.0	*1.5*	34.2	*6.0*	18.6	*0.8*	33.0	*7.8*	9.6	29.7	*6.5*
	cc-pVTZ[b]	10.87	*2.4*	32.42	*4.2*							
BLYP	6-31G(d)[b]	7.25	*-1.3*	23.75	*-4.5*							
	6-31++G(d,p)[b]	9.06	*0.6*	24.01	*-4.2*							
	cc-pVDZ[b]	7.59	*-0.9*	23.51	*-4.7*							
	6-311G(d,p)[b]	13.60	*5.1*	26.89	*-1.3*							
	cc-pVTZ[b]	14.09	*5.6*	26.98	*-1.2*							
G96LYP	6-31G(d)[b]	6.99	*-1.5*	24.20	*-4.0*							
	6-31++G(d,p)[b]	8.31	*-0.2*	24.18	*-4.0*							
	cc-pVDZ[b]	7.14	*-1.4*	23.83	*-4.4*							
	6-311G(d,p)[b]	12.98	*4.5*	26.98	*-1.2*							
	cc-pVTZ[b]	13.26	*4.8*	26.97	*-1.2*							
Experiment		8.5±1.0[c]		28.2±0.5[c]		17.8±1.2[d]		25.2±0.8[d]			23.2[e]	

[a] Ref. [24], geometries were optimized using the 6-31G(d,p) basis set
[b] Ref. [38], geometries were optimized using the 6-31G(d) basis set
[c] Ref. [39]
[d] Ref. [40]
[e] Ref. [41], using the relationship between E_a and H for an unimolecular reaction ($\Delta H^{\ddagger} = E_a - RT$)

3.3 Functional Effects

All calculations used two hybrid density functionals; the three parameter B3LYP functional and the one parameter PBE0 functional.[44] One of the advantages of the PBE0 functional is its ability to provide results superior to empirically derived density functionals for a variety of chemical properties (magnetic, vibrational and electronic).[45] This seems to be an obvious advantage over B3LYP, which has been shown to poorly predict magnetic and electronic properties for *p*-benzyne.[27]

Geometry optimizations performed with the functionals produced two sets of three geometries (Figure 5 and Supplemental Information). While the two methods generally reproduced each other's bond lengths within a hundredth of an angstrom, the non-bonded internuclear distances were significantly different. Figure 5 depicts the important non-bonded internuclear distances of the chemically active region for all six optimized geometries. In all cases, the PBE0 functional predicts shorter distances, and tighter angle bends within the ring. These differences are most significant in the pre-cyclized structure, where the PBE0 structure has the key carbon distances shortened by 0.024 and 0.026Å, respectively. This contraction effect is reduced in the transition state as the benzyne ring is partially formed, and the PBE0 structure has the same carbon distances shortened by 0.015 and 0.012Å relative to the B3LYP structure. Once the Bergman cyclization is completed and the *p*-benzyne ring is fully formed, the two different DFT structures now have carbon distances that differ by less than a hundredth of an angstrom.

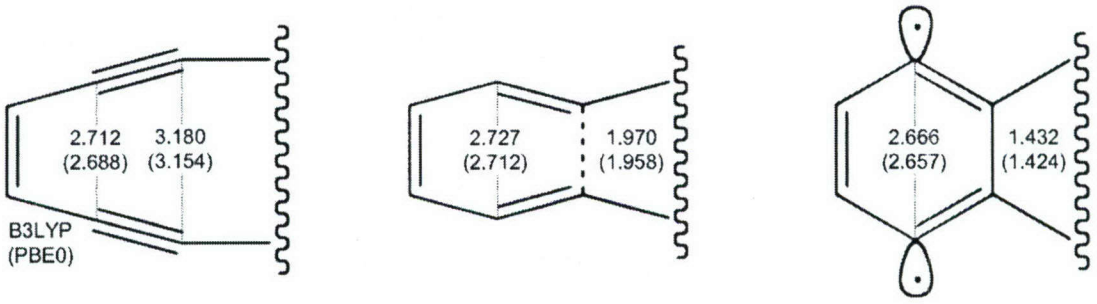


Figure 5. Critical internuclear distances for the three optimized MTC2 geometries. Non-bonded distances are displayed with a dotted line, the semi-bond of the transition state is displayed with a dashed line. Geometries were optimized with both the B3LYP and PBE0 (in parenthesis) functionals using the 6-31G(d,p) basis set. All values are given in angstroms.

These subtle differences in geometry appear to manifest themselves as significant differences to computed energy. While B3LYP/6-31G(d,p) predicts ΔE_e and $\Delta E_e^{\ddagger}$ values of -2.34 and 22.52 kcal•mol^{-1}, similar PBE0 calculations give dramatically different results; -11.93 and 19.62 kcal•mol^{-1}. Increasing the basis set size for subsequent single point calculations produces no significant changes to these results (as explained later). We believe that the large negative ΔE_e predicted by the PBE0 calculation stems, at least partially, from significant destabilization in the PBE0 optimized pre-cyclized MTC2 structure. Figure 6 illustrates this effect by depicting the B3LYP and PBE0 reaction coordinates. Because there were no significant geometric distortions between the two optimized diradical geometries, the two reaction coordinates were 'zeroed' energetically with respect to products. Once the product energies are normalized, the difference between the B3LYP and PBE0 results is apparent. The PBE0 reactant and transition-state structures are destabilized relative to those obtained with B3LYP.

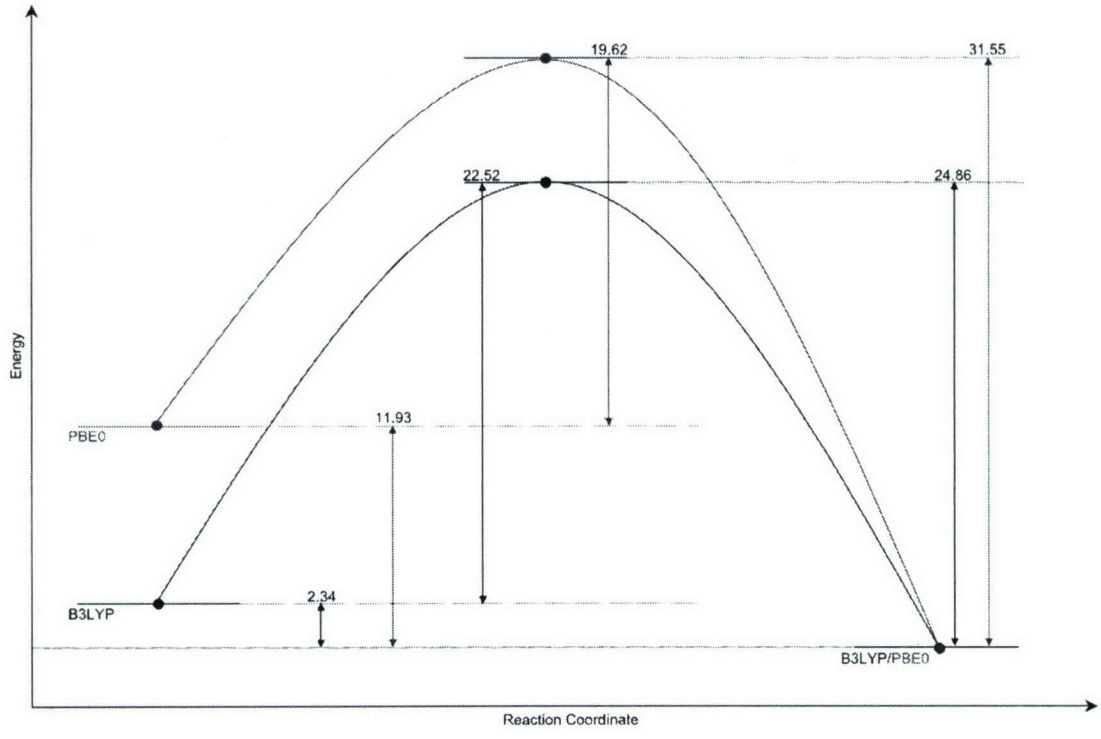


Figure 6. Reaction coordinate of the MTC2 Bergman cyclization. Both B3LYP (black) and PBE0 (red) reaction coordinates are displayed with both product species normalized to each other with respect to absolute energy. Relative electronic energies were calculated with the 6-31G(d,p) basis and are given in kcal•mol^{-1}.

3.4 Basis Set Effects

We used six different basis sets in this study in an effort to study which types of basis functions would be computationally efficient for larger enediyne systems. While all geometry optimizations used the well established 6-31G(d,p) basis set, a variety of other basis sets were tested using a series of single point calculations. Specifically, we tested the differences between similarly sized Pople and Dunning style basis sets, the importance of expanding from a double- to a triple-zeta basis set and what impact the augmentation of diffuse functions may have. Enthalpies, free energies and electronic energies of reaction calculated with these six basis sets are displayed in Table 2.

As was the case with the small enediynes (**1** and **2**) used for previous benchmarking studies,[24] an expansion of the basis set from double- to triple-zeta produced significant changes in the calculated energies of reaction. Predicted enthalpies of cyclization became ~5 kcal•mol^{-1} more exothermic for each of the three pairs of corresponding double-/triple-zeta basis sets. This trend was also present in the PBE0 data, as triple-zeta calculations were consistently ~4.5 kcal•mol^{-1} more exothermic than their double-zeta counter parts. These variations in enthalpies are slightly smaller than, but consistent with the variations in calculated enthalpies for **1** and **2**, which were ~7 kcal•mol^{-1} more exothermic for triple-zeta calculations, and tended to agree much better with experiment.[24] Enthalpies of reaction appeared to be less sensitive to this type of basis set expansion. While quantities still became more exothermic, the barrier heights to MTC2 cyclization only increased by ~3 – 3.5 and ~2.5 – 3 kcal•mol^{-1} for B3LYP and PBE0 respectively. This result is, again, very consistent with the results from the smaller benchmarking enediynes (**1**, **2** and **3**), as predicted barrier heights in these small molecules increased by ~2.5 – 3.5 kcal•mol^{-1} when the basis set was expanded from double- to triple-zeta.[24] The inclusion of triple-zeta single point calculations is essential for any computational schemes devised for this type of system.

Augmenting a basis set with diffuse functions had a negligible effect upon calculated energies for MTC2. In all four pairs of augmented/non-augmented calculations (two for B3LYP and two for PBE0) the enthalpies of reaction became more endothermic by less than 0.25 kcal•mol^{-1}. Similarly, changes to enthalpy barriers were also very small and endothermic, the largest change was less than 0.35 kcal•mol^{-1}. Diffuse functions are an unnecessary expense.

Differences between computations performed with the Pople and Dunning style basis sets are small as well, with the larger differences occurring at the triple-zeta basis level. For enthalpies of reaction, Dunning basis sets consistently produce predictions that are more exothermic than their Pople counterparts, with differences less than 0.35 kcal•mol^{-1} at the double-zeta and less than 0.90 kcal•mol^{-1} at the triple-zeta levels of theory. Changes to calculated barrier heights are small as well. Barrier heights at the double-zeta level increase by less than 0.30 kcal•mol^{-1} and by less than 0.50 kcal•mol^{-1} at the triple-zeta level. Again, these changes were consistent with the data from molecule **1**, where differences between calculations performed using the Pople and Dunning basis sets were quite minor. Also, the more exothermic Dunning style basis set calculations are in slightly better agreement with available experimental data for enediyne **1**. Furthermore, the Dunning basis sets are both smaller than their Pople counterparts and do not employ a segmented contraction scheme.[31] These two features make the Dunning basis sets more computationally efficient, and thus less expensive while producing slightly more accurate predictions than the Pople basis sets. Therefore, it would seem slightly advantageous to employ Dunning style basis sets for computations upon enediyne systems.

Table 2. Electronic energies of reaction and gas-phase DFT thermochemistry data as a function of basis set. All values presented in kcal•mol^{-1}.

	B3LYP/X//6-31G(d,p)					
	6-31G(d,p)	cc-pVDZ	aug-cc-pVDZ	6-311G(2d,p)	cc-pVTZ	aug-cc-pVTZ
ΔE_e	-2.34	-2.00	-1.76	2.67	3.57	3.52
$\Delta E_e^{\ddagger}$	22.52	22.62	22.58	25.72	26.23	26.10
ΔH	-2.50	-2.16	-1.92	2.51	3.41	3.36
$\Delta H^{\ddagger}$	20.60	20.70	20.65	23.80	24.30	24.18
ΔG	-1.03	-0.69	-0.45	3.98	4.88	4.83
$\Delta G^{\ddagger}$	21.55	21.65	21.61	24.75	25.26	25.13

	PBE0/X//6-31G(d,p)					
	6-31G(d,p)	cc-pVDZ	aug-cc-pVDZ	6-311G(2d,p)	cc-pVTZ	aug-cc-pVTZ
ΔE_e	-11.93	-11.70	-11.78	-8.01	-7.36	-7.38
$\Delta E_e^{\ddagger}$	19.62	19.90	19.57	22.12	22.43	22.30
ΔH	-11.93	-11.70	-11.78	-8.01	-7.36	-7.38
$\Delta H^{\ddagger}$	17.96	18.24	17.91	20.46	20.77	20.65
ΔG	-10.43	-10.20	-10.28	-6.52	-5.86	-5.88
$\Delta G^{\ddagger}$	18.93	19.21	18.88	21.43	21.74	21.61

3.5 Solvation Effects

A series of B3LYP/6-31G(d,p) geometry optimizations were performed in conjunction within a dielectric continuum solvation model. Optimized gas-phase structures were used as the input geometry. Converged aqueous-phase geometries were obtained for both the pre-cyclized and diradical species. Unfortunately the transition state geometry suffered from SCF convergence failures with the open shell wavefunction. Re-optimization of the pre-cyclized geometry produced one significant change to the chemically active ring structure, as the non-bonded carbon-carbon internuclear distance lengthened by 0.024Å. All bonded internuclear distances in the pre-cyclized region remained relatively constant ($\Delta r < 0.005$Å).

In addition to the aqueous geometry optimization calculations, a series of single point aqueous calculations were performed on optimized gas-phase geometries using the Gaussian SCRF option, SCFVAC. The aqueous-phase geometries yielded electronic energies of reaction of -1.23 and 3.79 kcal•mol^{-1} for the 6-31G(d,p) and 6-311G(2d,p) basis sets respectively, whereas the SCFVAC calculations yielded electronic energies of

reaction of -2.87 and 2.17 kcal•mol^{-1} for the two basis sets, a change of ~1.6 kcal•mol^{-1}. The aqueous-phase free energies of reaction and activation predicted by SCFVAC for MTC2 are 2.08 and 24.03 kcal•mol^{-1} respectively, calculated at the triple-zeta level of theory. These results suggest that the computationally expensive, and labor intensive process of geometry re-optimization in PCM may be unnecessary to achieve accurate aqueous phase reaction data.

4. Conclusion

We have established an effective computational methodology for the study of Bergman cyclization in small and moderate sized enediyne molecules. The shortcomings of the PBE0 results underscore the need for continued development of density functionals specifically designed to study biomolecules, particularly those with radical and diradical chemistry. While the venerable B3LYP functional is the most cost effective way to study the energetics of the Bergman cyclization, numerous questions about B3LYP's ability to accurately model spin, and its associated magnetic properties still persist. Also, accurately modeling the two open shell states (transition state and diradical) is challenging. BS-UDFT does a good job of generating open shell singlet states, however it does fail at times, so spin excess values must be rigorously checked for wavefunction collapse. Stability calculations provide dependable alternatives to BS-UDFT, these calculations however are very expensive for any system much larger than MTC2 and prohibitively expensive for esperamicin A_1. A cheaper, yet effective, alternative is needed. Repeating a calculation with combinations of symmetry breaking and valence orbital shifting, combined with checking $\langle S^2 \rangle$ values, is the most cost effective procedure

at present. We will employ our methodology in the future study of larger enediyne model systems, and the eventual study of esperamicin A_1.

Acknowledgements

Acknowledgment is made to the donors of The Petroleum Research Fund, administered by the ACS, to Research Corporation, to NIH, to DOD, and to Hamilton College for support of this work. This project was supported in part by NSF Grant CHE-0457275, and by NSF grants CHE-0116435 and CHE-0521063 as part of the MERCURY high-performance computer consortium (http://mercury.chem.hamilton.edu).

Supporting Information Available:

Coordinates, electronic energies, thermochemistry corrections, and solvation energies for the precyclized MTC2 and transition state structures. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Konishi, M.; Ohkuma, H.; Saitoh, K. I.; Kawaguchi, H.; Golik, J.; Dubay, G.; Groenewold, G.; Krishnan, B.; Doyle, T. W. *Journal of Antibiotics* 1985, *38*, 1605.

(2) Golik, J.; Dubay, G.; Groenewold, G.; Kawaguchi, H.; Konishi, M.; Krishnan, B.; Ohkuma, H.; Saitoh, K.; Doyle, T. W. *J. Am. Chem. Soc.* 1987, *109*, 3462.

(3) Golik, J.; Clardy, J.; Dubay, G.; Groenewold, G.; Kawaguchi, H.; Konishi, M.; Krishnan, B.; Ohkuma, H.; Saitoh, K.; Doyle, T. W. *J. Am. Chem. Soc.* 1987, *109*, 3461.

(4) Bergman, R. G. *J. Am. Chem. Soc.* 1973, *6*, 25.

(5) Long, B. H.; Golik, J.; Forenza, S.; Ward, B.; Rehfuss, R.; Dabrowiak, J. C.; Catino, J. J.; Musial, S. T.; Brookshire, K. W.; Doyle, T. W. *Proceedings of the National Academy of Sciences of the United States of America* 1989, *86*, 2.

(6) Lode, H. N.; Reisfeld, R. A.; Handgretinger, R.; Nicolaou, K. C.; Gaedicke, G.; Wrasidlo, W. *Cancer Research* 1998, *58*, 2925.

(7) Nicolaou, K. C.; Dai, W. M. *Angewandte Chemie-International Edition* 1991, *30*, 1387.

(8) Nicolaou, K. C.; Hong, Y. P.; Torisawa, Y.; Tsay, S. C.; Dai, W. M. *J. Am. Chem. Soc.* 1991, *113*, 9878.

(9) Christner, D. F.; Frank, B. L.; Kozarich, J. W.; Stubbe, J.; Golik, J.; Doyle, T. W.; Rosenberg, I. E.; Krishnan, B. *J. Am. Chem. Soc.* 1992, *114*, 8763.

(10) Ikemoto, N.; Kumar, R. A.; Dedon, P. C.; Danishefsky, S. J.; Patel, D. J. *J. Am. Chem. Soc.* 1994, *116*, 9387.

(11) Smith, A. L.; Nicolaou, K. C. *Journal of Medicinal Chemistry* 1996, *39*, 2103.

(12) Cramer, C. J. *J. Am. Chem. Soc.* 1998, *120*, 6261.

(13) Johnson, W. T. G.; Cramer, C. J. *Journal of Physical Organic Chemistry* 2001, *14*, 597.

(14) Johnson, W. T. G.; Cramer, C. J. *J. Am. Chem. Soc.* 2001, *123*, 923.

(15) Alabugin, I. V.; Manoharan, M. *J. Phys. Chem. A* 2003, *107*, 3363.

(16) Alabugin, I. V.; Manoharan, M. *J. Am. Chem. Soc.* 2003, *125*, 4495.

(17) Tuttle, T.; Kraka, E.; Cremer, D. *J. Am. Chem. Soc.* 2005, *127*, 9469.

(18) Pickard, F. C.; Shepherd, R. L.; Gillis, A. E.; Dunn, M. E.; Feldgus, S.; Kirschner, K. N.; Shields, G. C.; Manoharan, M.; Alabugin, I. V. *J. Phys. Chem. A* 2006, *110*, 2517.

(19) *Anticancer agents from natural products*; Cragg, G. M.; Kingston, D. G. J.; Newman, D. J., Eds.; CRC Press: Boca Raton, FL, 2005, pp 577.

(20) Magnus, P.; Fortt, S.; Pitterna, T.; Snyder, J. P. *J. Am. Chem. Soc.* 1990, *112*, 4986.

(21) Snyder, J. P. *J. Am. Chem. Soc.* 1990, *112*, 5367.

(22) Lindh, R.; Ryde, U.; Schutz, M. *Theoretical Chemistry Accounts* 1997, *97*, 203.

(23) Zeidan, T. A.; Manoharan, M.; Alabugin, I. V. *Journal of Organic Chemistry* 2006, *71*, 954.

(24) Feldgus, S.; Shields, G. C. *Chem. Phys. Lett.* 2001, *347*, 505.

(25) Crawford, T. D.; Kraka, E.; Stanton, J. F.; Cremer, D. *J. Chem. Phys.* 2001, *114*, 10638.

(26) Brummel, H. A.; Shields, G. C. *Int. J. Quantum Chem.* 1995, *56*, 51.

(27) Grafenstein, J.; Kraka, E.; Filatov, M.; Cremer, D. *Int. J. Mol. Sci.* 2002, *3*, 360.

(28) Krishnan, B.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J. Chem. Phys.* 1980, *72*, 650.

(29) McLean, A. D.; Chandler, G. S. *J. Chem. Phys.* 1980, *72*, 5639.

(30) Dunning, T. H. *J. Chem. Phys.* 1989, *90*, 1007.

(31) Cramer, C. J. *Essentials of Computational Chemistry: Theories and Models*, 2nd ed.; Wiley, 2004.

(32) Cances, E.; Mennucci, B.; Tomasi, J. *J. Chem. Phys.* 1997, *107*, 3032.

(33) Mennucci, B.; Tomasi, J. *J. Chem. Phys.* 1997, *106*, 5151.

(34) Cossi, M.; Barone, V.; Mennucci, B.; Tomasi, J. *Chem. Phys. Lett.* 1998, *286*, 253.

(35) Frisch, M. J. T., G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, Jr., J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.;

Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. Gaussian 03; Revision C.02 ed.; Gaussian, Inc.: Wallingford, CT, 2004.

(36) Grafenstein, J.; Cremer, D. *Molecular Physics* 2001, *99*, 981.

(37) Bauernschmitt, R.; Ahlrichs, R. *J. Chem. Phys.* 1996, *104*, 9047.

(38) Prall, M.; Wittkopp, A.; Schreiner, P. R. *J. Phys. Chem. A* 2001, *105*, 9265.

(39) Roth, W. R.; Hopf, H.; Horn, C. *Chemische Berichte* 1994, *127*, 1765.

(40) Roth, W. R.; Hopf, H.; Wasser, T.; Zimmermann, H.; Werner, C. *Liebigs Annalen* 1996, 1691.

(41) Nicolaou, K. C.; Zuccarello, G.; Ogawa, Y.; Schweiger, E. J.; Kumazawa, T. *J. Am. Chem. Soc.* 1988, *110*, 4866.

(42) Kraka, E.; Cremer, D. *J. Am. Chem. Soc.* 2000, *122*, 8245.

(43) Grafenstein, J.; Hjerpe, A. M.; Kraka, E.; Cremer, D. *J. Phys. Chem. A* 2000, *104*, 1748.

(44) Liptak, M. D.; Shields, G. C. *Int. J. Quantum Chem.* 2005, *105*, 580.

(45) Adamo, C.; Barone, V. *J. Chem. Phys.* 1999, *110*, 6158.

Computational Design and Experimental Discovery of an Anti-estrogenic Peptide Derived from Alpha-Fetoprotein

Karl N. Kirschner,[a]* Katrina W. Lexa,[a] Amanda M. Salisburg,[a] Katherine A. Alser, [a] Leroy Joseph,[b] Thomas T. Andersen, [b]* James A. Bennett,[c] Herbert I. Jacobson,[d] and George C. Shields[a]*

[a] Hamilton College
Department of Chemistry
Center for Molecular Design
198 College Hill Road
Clinton, NY 13323

[b] Albany Medical College
Center for Cardiovascular Sciences
Albany, NY 12208

[c] Albany Medical College
Center for Immunology and Microbial Disease
Albany, NY 12208

[d] Albany Medical College
Department of Obstetrics, Gynecology, and Reproductive Science
Albany, NY 12208

Abstract

State-of-the-art molecular dynamics simulations predict a common conformation shared between 8-mer and cyclic 9-mer peptides that inhibit human breast cancer. The predicted common structure contains a conserved reverse turn. The simulations were used to identify the smaller peptide analogs TOVNO, TPVNP, TOVN, and TPVN with the same conserved reverse turn. These analogs have been synthesized and shown to inhibit estrogen-dependent cell growth in a mouse uterine growth assay, a test that shows reliable correlation with the inhibition of human breast cancer. This is the first time that Replica Exchange Molecular Dynamics has been used to predict a new lead drug

compound.

Introduction

Breast cancer is the most common cancer diagnosed in women and is the second leading cause of cancer death among women, closely following lung cancer. In 2006, the American Cancer Society estimated that 212,920 women in the United States will be diagnosed with invasive breast cancer and predicted 41,430 deaths (*1*). Tamoxifen is the most widely used drug for the treatment of estrogen receptor-positive breast cancer, acting through competition with estrogen for binding to the estrogen receptor (ER).

However, approximately one third to one sixth of ER-positive breast cancers are intrinsically resistant to tamoxifen, and many more acquire resistance to this drug during treatment (*2*). Additionally, tamoxifen stimulates uterine growth, which can lead to uterine cancer in a small percentage of women taking this drug (*2*). Consequently, there is ongoing need for breast cancer drugs with greater efficacy and fewer side effects.

Alpha-fetoprotein (AFP) is an embryo specific serum alpha-globulin glycoprotein that is synthesized primarily by the fetal liver and circulates through the serum of pregnant women (*3*). AFP has been reported to bind and transport ligands, including fatty acids, steroids, heavy metal ions, phytoestrogens, drugs and some toxins (*4*). AFP is a growth-regulating hormone with the capacity to stimulate or inhibit growth (*5*). From fertilization through birth, AFP holds an important physiological role as a developmental promoter for the fetus. More recent reports have shown that AFP has antiestrogenic activity and can inhibit the growth of estrogen dependent cancer (*2, 6-8*). These data, combined with epidemiological data showing that elevated levels of AFP are associated

with reduced lifetime risk of breast cancer (*7*), have led to the suggestion that AFP or analogs thereof may be a useful agent for chemoprevention, as well as for the treatment, of breast cancer (*7, 9, 10*). Festin et al. (*11*), Eisele et al. (*12, 13*), and Mesfin et al. (*14*) parsed the 70,000 MW AFP into to a series of smaller peptides that retained the same anti-estrogenic and anti-breast cancer activities of the original protein. Their work resulted in the smallest known active analog, an 8-mer with sequence EMTPVNPG (AFPep) (*14*), which consists of amino acids 472-479 from the human AFP sequence. Amino acid substitution studies revealed the importance of specific residues essential for activity (*15*). All efforts since then to create an active peptide consisting of fewer than eight residues has resulted in the loss of anti-estrogenic activity (*14, 15*).

A receptor for AFP on cancer cell membranes has been reported (*16*), and binding of AFP (or AFPep) to this receptor signals the cell to inhibit its own growth. The steric/electronic features of the receptor site that permits AFP binding are largely unknown, making rational development of lead compounds difficult. We present here a novel strategy for developing new lead compounds, a strategy that uses molecular dynamics to explore the allowed conformational space of potentially active analogs in solution. Rational drug design using molecular dynamics in this manner involves understanding the conformational space occupied by the active compounds followed by the creation of different compounds that sample the same space. We have used Replica Exchange (*17, 18*) Molecular Dynamics (REMD) techniques (*19*) to explore the conformational dynamics of several AFPep analogs. Our REMD results reveal that the peptide's critical region for activity is a four amino acid sequence that adopts a Type I β-turn conformation. We have run REMD simulations on several different four- and five-

amino acid peptides, synthesized those that appeared promising, along with controls, and tested them for activity using an immature mouse uterine growth assay (*2*). Results from the REMD simulations and the experimental activity studies are presented in this paper, and show that peptides as short as four amino acid residues retain biological activity.

The REMD technique has been successfully used to obtain energy landscapes for the TRP cage and for the C-terminal β-hairpin of protein G (*20-22*), to find the global minimum for chignolin (*23*), and to explore unfolding of alpha helical peptides as a function of pressure and temperature (*24*). REMD has also been used to explore intrafacial folding and membrane insertion of designed peptides (*25*), and protein folding pathways (*26, 27*). We demonstrate for the first time that REMD, besides being an excellent technique for probing peptide and protein structure, can be used for drug design.

Results

Figure 1 shows the conservation of conformational space of the five active 8-mer and cyclic 9-mer AFP-derived peptides previously shown (15) to have anti-estrogenic activity. This figure shows the overlay of the most representative peptide geometries obtained from the conformational family that displays the common reverse turn motif in each REMD simulation. Four amino acids, TPVN, are conserved in the conformational space of the active 8-mers and cyclic 9-mers. The TPVN sequence forms a reverse turn within the longer peptides, a structure that is conserved across all five REMD simulations. In a Type I reverse β-turn, four amino acids form a turn structure, which is defined by the phi (ϕ) and psi (ψ) angles of the second and third amino acids. An ideal

Type I β-turn has ϕ/ψ values of -60 and -30 degrees for the second amino acid, and ϕ/ψ values of -90 and 0 degrees for the third amino acid in the four amino acid turn sequence. As displayed in Table 1, the cyclic-[EMTPVNPGQ] and EMTPVNPG peptides have average ϕ_{Pro}, ψ_{Pro}, and ϕ_{Val}, ψ_{Val} values that fall within 14 degrees of a Type I turn. The other three simulations of the 8-mer and 9-mer peptides have similar values (Figure 1).

We ran REMD simulations for the TPVN, TPVNP, and PVNP analogs to explore the conformational space sampled by these four and five amino acid peptides. The TPVN and TPVNP peptides form the same reverse turn seen in the larger, active peptides. As displayed in Table 1, the TPVNP and TPVN analogs have ϕ_{Pro}, ψ_{Pro} values of approximately -67 and -19 degrees, while the ϕ_{Val}, ψ_{Val} values are approximately -91 and -12 degrees. In contrast, the average ϕ_{Pro}, ψ_{Pro}, and ϕ_{Val}, ψ_{Val} values for the three different conformations of the PVNP peptide reveal that none of the three structures is a turn (Table 1). The structure that is sampled for 41% of the REMD simulations has ϕ_{Pro}, ψ_{Pro} angles of -75 and –14 degrees, and ϕ_{Val}, ψ_{Val} angles of –109 and 120 degrees. The structure sampled for 37% of the REMD simulation has ϕ_{Pro}/ψ_{Pro} of –108 and 1 degrees, while the structure sampled for 27% of the time has ϕ_{Pro}/ψ_{Pro} of –107 and 46 degrees. All three of these structures have ϕ_{Val}/ψ_{Val} of –126 and 123 degrees. Consideration of the average valine angles alone show that the PVNP peptide does not form a turn structure.

The representative dynamics of three analogs containing the conserved TPVN sequence can be visualized in Figure 2. The top graphs display the distance between the C_α atoms of the threonine and asparagine residues, serving as a definitive diagnostic for a β-turn. Twenty-five percent of β-turns have no intraturn hydrogen bond, so an alternative definition of a β-turn is that the distance between the C_α carbon atoms of amino acid

residues one and four in the tetrapeptide sequence is less than 7 Å (*28*). The TPVN tetrapeptide has a $C_{\alpha(T)} - C_{\alpha(N)}$ distance less than 7 Å for 64% of the simulation, indicating a β-turn. The TPVNP pentapeptide adopts the turn structure for 75% of the simulation, and the cyclic-[EMTPVNPGQ] peptide is locked into the turn structure for 99% of the simulation. The bottom graphs in Figure 2 show the corresponding three-dimensional plots of ϕ versus ψ versus simulation time. These plots reveal the dynamics of the ϕ/ψ values for proline (red) and valine (green) amino acids throughout the simulation for these three peptides. These plots confirm that these three peptides adopt a β-turn conformation over the course of the simulation.

Several 4-mer and 5-mer peptide analogs containing the TPVN sequence, or a similar sequence with hydroxyproline (O) substituted for proline, were synthesized, tested, and compared to the original 8-mer peptides for biological activity. As shown in Table 2, biological activity, as defined by inhibition of estrogen-stimulated growth of immature mouse uterus, was retained in the TPVN and TOVN 4-mers and even more so in the TPVNP and TOVNO 5-mers.

The OVNO and PVNP analogs, which were thought to represent the pharmacophore region of AFPep (*15*), did not show significant activity. Similarly, five amino acid peptides containing the amino or carboxyl end of AFPep did not have biological activity (bottom of Table 2). The 7-mer, EMTPVNP, was slightly less active than AFPep and the smaller TPVNP.

The figures reveal why the TPVNP and TPVN analogs are active. All active peptides have a conserved reverse β-turn motif. These β-turns are formed by the TPVN sequence, with a hydrogen bond formed between the carbonyl oxygen of the first amino

acid and the amide hydrogen of the fourth amino acid. The conservation of proline in the second position favors the formation of a reverse turn. This proline is conserved in human, gorilla, chimpanzee, horse, rat, and mouse AFP sequences (*15*).

Figure 1. Overlay of the cyclic-[EKTPVNPGN] (red), cyclic-[EKTPVNPGQ] (blue), cyclic-[EMTPVNPGQ] (orange), and the EMTPVNPG (green) and EMTPTNPG (purple) peptides from REMD simulations. Each structure represents the β-turn motif sampled during the dynamics.

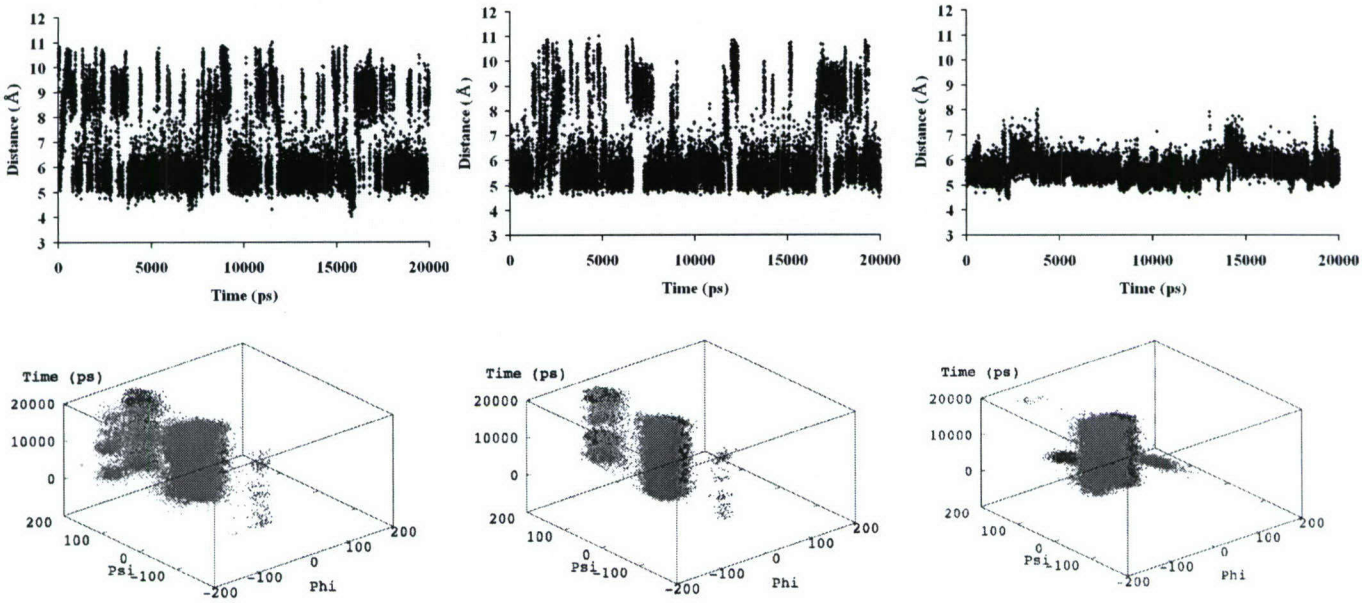


Figure 2. The top graphs depict the distances between the conserved threonine and asparagine C_α atoms as a function of simulation time for the TPVN (left), TPVNP (middle), and cyclic-[EMTPVNPGQ] peptides. The bottom graphs depict their corresponding phi (x-axis) and psi (y-axis) angles as a function of simulation time for the conserved proline (red) and valine (green) amino acids.

Table 1. Average ϕ/ψ angle values for the second and third amino acids for five different sequences. The four residues in boldface define amino acids one through four for each peptide. The ϕ/ψ angle values for the second and third amino acids of a four amino acid sequence are diagnostic for a β-turn structure.

		ϕ_2	ψ_2	ϕ_3	ψ_3
Type I β-turn		-60	-30	-90	0
Cyclic-[EM**TPVN**PGQ]		-68	-16	-96	-9
EM**TPVN**PG[a]		-68	-19	-92	-11
TPVNP		-66	-19	-94	-12
TPVN		-68	-19	-88	-11
PVNP	41%	-75	-14	-126	123
	32%	-108	1	-126	123
	27%	-107	46	-126	123

a. The sequence of the AFPep peptide.

Table 2. Effect of AFP-derived peptides on estrogen-stimulated growth of immature mouse uterus. Estradiol and peptide injection into immature mice was carried out as described in Material and Methods. Peptide was injected intraperitoneally at a dose of 1μg per mouse.

Sequence	% Inhibition ± SE
EMTPVNPG	34 ± 3[a]
EMTOVNOG	32 ± 4[a]
EMTPVNP	19 ± 2
TOVNO	31 ± 4[a]
TPVNP	26 ± 1[a]
TPVN	24 ± 5[a]
TOVN	22 ± 1[a]
OVNO	6 ± 5
PVNP	5 ± 3
PGVGQ	0
EMTPV	0
EMTOV	0

a. $p < 0.05$ when compared to group stimulated with E_2 alone. Wilcoxon Rank-Sum Test.

Discussion

We have used REMD simulations to sample the conformational space of 8-mer and 9-mer AFP-derived peptides that have anti-estrogenic and anti-breast cancer activity. We discovered that an identical four amino acid sequence had minimal conformational flexibility, suggesting that this region is essential for the biological activities of these peptides. The TPVN, TOVN, TPVNP, and TOVNO sequences were subsequently synthesized and were found to be biologically active. This is a novel finding because Mesfin et al.(*11*) and DeFreest et al.(*15*) had concluded that the 8-mer, EMTPVNPG, was the minimum number of residues that an AFP-derived peptide can have and still retain significant anti-estrogenic activity. This conclusion was based on the findings that the 7-mers, MTPVNPG and EMTPVNP, had relatively less biological activity than the 8-mers. These investigators concluded that the 8-mer peptide assumed a pseudo-cyclic conformation, that the middle of the peptide contained its pharmacophore region (the PVNP sequence), and that the exterior residues were essential for holding the peptide in its active conformation. The 7-mers lost activity, which was assumed to result from the loss of hydrogen bonding between the first and eighth amino acid residues, suggesting that full activity must require at least eight amino acid residues. Indeed the 5-mer, EMTPV, was synthesized and found to have no significant activity, which supported their conclusion that the 8-mer sequence was the minimal sequence for biological activity (*14*). Furthermore, guided by this hydrogen bond hypothesis, cyclic 9-mers were also synthesized and shown to possess significant biological activity (*29*). We tested this 8-mer hypothesis by examining the percentage of hydrogen bond formation between the first and eighth amino acids for each simulation containing seven or eight amino acid

residues. For all 8-mer and 7-mer simulations, hydrogen bond formation between the first and the last amino acids is observed for less than 1% of the simulation if the distance definition for a hydrogen bond is set at 2.3 Å. This result is contrary to the original hypothesis.

The REMD studies reported herein lead to a substantially different conclusion regarding the requirement for peptide activity. These studies indicate that the key region is the TPVN sequence, since this structure retains the turn conformation common in every active peptide; the PVNP peptide does not sample a reverse turn structure (Table 1) and shows relatively low activity (Table 2). The above discovery has significant consequences. First, a 4-mer or a 5-mer peptide that retains biological activity is less expensive to synthesize than an 8-mer or 9-mer peptide, and these smaller analogs are more drug-like. Second, knowing the conformational space occupied by the 4-mer provides insight into the topology of the unknown receptor site for these peptides. Based on this study, the receptor site topology for the AFP analogue peptides is predicted to be a mirror image of a reverse β-turn. Previous substitutions in the 8-mer sequences reveal that when threonine, leucine, or isoleucine is substituted for valine, biological activity is retained, while substitution of valine with D-valine or alanine results in loss of activity. This implies that the topology of the receptor is stereospecific, and branched amino acids are essential for creation of hydrophobic forces that bind the receptor to the peptide. Finally, it leads to a different conceptual approach for stabilization of these peptides and the development of peptidomimetics. Peptidomimetics can be designed based on the active 4-mer and 5-mer peptides, with REMD used to ensure that the steric and electronic nature of the peptides is retained. Developing a peptidomimetic may not be necessary, for

as long as the peptides are bioavailable they have the advantage of a lower probability of side effects compared to peptide mimics. The AFP 8-mer and cyclic 9-mer peptides have already been shown to be non-toxic and bioavailable in mice (*2*). Cancer xenograft assays in mice involved the administration of 8-mer and cyclic 9-mer peptides twice a day for 30 days, during which there was no change in mouse body weight, cage activity, fur texture, or body temperature. In addition there were no changes in size or appearance of major organs relative to the control group. The uterine growth studies revealed that these peptides, unlike tamoxifen, did not stimulate murine uterine growth; indeed they inhibit the uterine stimulated growth induced by tamoxifen (*2, 8*). Thus, the peptides derived from AFP represent a new class of potential breast cancer drugs, which are active through a new, yet to be discovered, receptor.

Because there is excellent correlation between the uterine growth assay and the human breast cancer xenograft assay with regard to AFPep peptide inhibition of estrogen-stimulated growth (*2, 14*), the TOVNO, TPVNP, TOVN, and TPVN analogs that are all active in the uterine growth assay are predicted to inhibit human breast cancer growth. We have begun the xenograft assays on these analogs to evaluate this prediction.

Conclusion

We have applied and demonstrated for the first time that REMD simulations can be used as a novel drug design tool. We have shown that REMD predicts a common conformation that is shared between the active linear 8-mer and cyclic 9-mer peptides. The predicted common conformation is a conserved reverse β-turn, and the smaller peptide analogs TOVNO, TPVNP, TOVN, and TPVN also contain the same conserved

reverse turn. These analogs are shown to inhibit estrogen-dependent cell growth in a mouse uterine growth assay, through interaction with a yet to be discovered key receptor, and are predicted to inhibit human breast cancer. The 5-mer and 4-mer peptides are new discoveries that may lead to promising new anti-breast cancer drugs (*30*).

References

1. **"Breast Cancer Facts & Figures 2005-2006," American Cancer Society, Atlanta, http://www.cancer.org/docroot/STT/stt_0.asp, June 15, 2006.**
2. **J. A. Bennett, F. B. Mesfin, T. T. Andersen, J. F. Gierthy, H. I. Jacobson, *Proceedings of the National Academy of Sciences of the United States of America* 99, 2211 (Feb 19, 2002).**
3. **G. I. Abelev, *Advances in Cancer Research* 14, 295 (1971).**
4. **G. J. Mizejewski, *Experimental Biology and Medicine* 226, 377 (May, 2001).**
5. **E. Dudich *et al.*, *Tumor Biology* 19, 30 (Jan-Feb, 1998).**
6. **H. I. Jacobson, J. A. Bennett, G. J. Mizejewski, *Cancer Research* 50, 415 (Jan 15, 1990).**
7. **J. A. Bennett, S. J. Zhu, A. Pagano-Mirarchi, T. A. Kellom, H. I. Jacobson, *Clinical Cancer Research* 4, 2877 (Nov, 1998).**
8. **R. R. Parikh *et al.*, *Clinical Cancer Research* 11, 8512 (Dec 1, 2005).**
9. **H. I. Jacobson, D. T. Janerich, *Biological Activities of Alpha-Fetoprotein.* G. J. Mizejewski, H. I. Jacobson, Eds. (CRC, Boca Raton, FL, 1989), pp. 94-100.**
10. **B. E. Richardson *et al.*, *American Journal of Epidemiology* 148, 719 (1998).**
11. **S. M. Festin *et al.*, *Biochimica Et Biophysica Acta-General Subjects* 1427, 307 (Apr 19, 1999).**
12. **L. E. Eisele *et al.*, *Journal of Peptide Research* 57, 539 (Jun, 2001).**
13. **L. E. Eisele *et al.*, *Journal of Peptide Research* 57, 29 (Jan, 2001).**
14. **F. B. Mesfin, J. A. Bennett, H. I. Jacobson, S. J. Zhu, T. T. Andersen, *Biochimica Et Biophysica Acta-Molecular Basis of Disease* 1501, 33 (Apr 15, 2000).**
15. **L. A. DeFreest *et al.*, *Journal of Peptide Research* 63, 409 (May, 2004).**
16. **M. J. Villacampa *et al.*, *Biochemical and Biophysical Research Communications* 122, 1322 (1984).**
17. **Y. Sugita, Y. Okamoto, *Chemical Physics Letters* 314, 141 (Nov 26, 1999).**
18. **A. Mitsutake, Y. Sugita, Y. Okamoto, *Biopolymers* 60, 96 (2001).**
19. **D. A. Case *et al.* (University of California, San Francisco, 2004) pp. AMBER Version 8 Molecular Dynamics Program.**
20. **R. H. Zhou, B. J. Berne, R. Germain, *Proceedings of the National Academy of Sciences of the United States of America* 98, 14931 (Dec 18, 2001).**
21. **R. H. Zhou, *Proteins-Structure Function and Genetics* 53, 148 (Nov 1, 2003).**
22. **R. H. Zhou, *Proceedings of the National Academy of Sciences of the United***

States of America **100, 13280 (Nov 11, 2003).**
23. **M. M. Seibert, A. Patriksson, B. Hess, D. van der Spoel, *Journal of Molecular Biology* 354, 173 (Nov 18, 2005).**
24. **D. Paschek, S. Gnanakaran, A. E. Garcia, *Proceedings of the National Academy of Sciences of the United States of America* 102, 6765 (May 10, 2005).**
25. **W. Im, C. L. Brooks, *Proceedings of the National Academy of Sciences of the United States of America* 102, 6771 (May 10, 2005).**
26. **H. Nymeyer, S. Gnanakaran, A. E. Garcia, in *Numerical Computer Methods, Pt D.* (2004), vol. 383, pp. 119.**
27. **M. Andrec, A. K. Felts, E. Gallicchio, R. M. Levy, *Proceedings of the National Academy of Sciences of the United States of America* 102, 6801 (May 10, 2005).**
28. **K.-C. Chou, *Analytical Biochemistry* 286, 1 (2000/11/1, 2000).**
29. **F. B. Mesfin, T. T. Andersen, H. I. Jacobson, S. Zhu, J. A. Bennett, *Journal of Peptide Research* 58, 246 (Sep, 2001).**
30. We thank Carlos Simmerling and Adrian Roitberg for insightful discussion of the replica exchange methodology, and Steve Festin for his contributions. Acknowledgment is made to the donors of the Petroleum Research Fund, administered by the ACS, to Research Corporation, to NIH (R01 CA102540 and R15 CA115524), to DOD (BC031067 and W81XWH-05-1-0441), and to Hamilton College for support of this work. This project was supported in part by NSF Grant CHE-0457275, and by NSF grants CHE-0116435 and CHE-0521063 as part of the MERCURY high-performance computer consortium (http://mercury.chem.hamilton.edu). This work was first presented at the 2006 International Symposium on Theory and Computations in Molecular and Materials Sciences, Biology and Pharmacology, on February 26, 2006.

SUPPORTING ONLINE MATERIAL

Methods

We used the AMBER 8 molecular dynamics program package (*19, 31*) with a new force field that has been created specifically to improve the representation of peptide structure (*32*). Simulations were run in implicit solvent, using the Generalized Born (GB) model (*33*) implemented in AMBER (*34*) with the default radii under the multi*sander* framework (*34*). It has been demonstrated that the use of continuum solvent models in REMD can lead to over stabilization of ion pairs that affect the secondary structure (*35, 36*), but since we do not have salt bridges in our peptide structure we avoid this potential

problem. Adequate sampling of conformational space was insured through the use of the Replica Exchange algorithm (*17, 18*). REMD was developed as a means to overcome local potential energy barriers. Compiled under AMBER 8, one-dimensional REMD explores a generalized canonical ensemble of N non-interacting replicas that undergo simulation separately but concomitantly at exponentially related temperatures, with exchanges between replicas occurring at a specified time interval. The consequence of this exchange is that entrapment in local potential energy wells is avoided. Those replicas, which do become trapped within a local well at one temperature, can escape when transitioned to a higher temperature as part of the exchange process. Thus, accuracy of conformation is maintained through analysis at low temperatures while simulations at higher temperatures efficiently achieve exploration of the potential energy surface.

The sequences used for the simulations were chosen from the set of previously synthesized active analogs (*15*). REMD simulations were run on the cyclic analogs cyclic-[EKTPVNPGN], cyclic-[EKTPVNPGQ], cyclic-[EMTPVNPGQ], and the linear analogs EMTPVNPG and EMTPTNPG. In addition, REMD simulations were run on the smaller analogs EMTPVNP, MTPVNPG, TPVNP, TPVN, and PVNP. Eight different replicas were used, each defined initially with the same input structure. Temperatures were selected to agree with exponential growth such that interchange occurred within the temperature group $temp_i$, $temp_o$= 265K, 304K, 350K, 402K, 462K, 531K, 610K and 700K. The number of exchange attempts between neighboring replicas was set to 1000 and the number of MD steps between exchange attempts was defined as 10000. Thus, the total length for each simulation with a time step of 0.002 picoseconds was 20 nanoseconds (*19*).

The linear peptide analogs shown in Table 2 were prepared using Fmoc solid-phase synthesis (*14, 15, 29*). The anti-estrogenic activity of each peptide was determined using the immature mouse uterine growth assay as described previously (*2, 7*). Intraperitoneal (i.p.) administration of 0.5 μg of 17β-estradiol (E_2) to an immature female mouse doubles its uterine weight in 24 hours (*7*). Swiss/Webster female mice, weighing 6-8 grams at 13-15 days old, were obtained from Taconic Farms. Mice were grouped so that each group had the same range of body weights. Each group received two sequential i.p. injections spaced one hour apart. The peptide or a saline solution control was contained in the first injectant, and E_2 or a saline control was contained in the second injectant. Twenty-two hours after the second injection, uteri were dissected and weighed immediately. The uterine weights were normalized to mouse body weights (mg uterine/g body). Experiments used a minimum of five mice per group, and the mean normalized uterine weight and standard deviations were determined for each group. Percent growth inhibition in a test group was calculated from the normalized uterine weights as given by equation 1.

$$\text{Growth inhibition (\%)} = (\text{Full } E_2 \text{ stimulation} - E_2 \text{ stimulation in test group})/(\text{Full } E_2 \text{ stimulation} - \text{No } E_2 \text{ stimulation}) \times 100\% \qquad \mathbf{1}$$

Significance of differences between groups was evaluated with the nonparametric Wilcoxon ranks sum test (one-sided). Generally, drug-induced growth inhibitions of 20% or greater are significantly different ($p \leq 0.05$) from the group receiving E_2 alone. Each AFPep analog was evaluated for antiuterotrophic activity in three or more experiments, and the mean growth inhibition ± the standard error for each analog is reported in Table 2.

Additional References for Methods

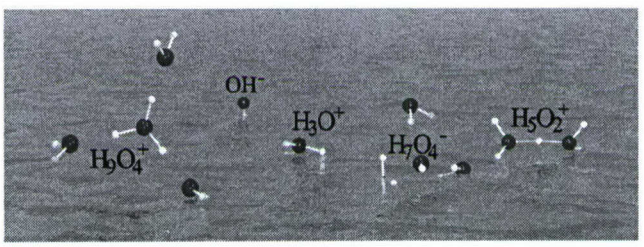
31. D. A. Case *et al.*, *Journal of Computational Chemistry* 26, 1668 (Dec, 2005).
32. V. Hornak *et al.*, *Proteins-Structure Function and Bioinformatics* in press (2006).
33. C. J. Cramer, *Essentials of Computational Chemistry: Theories and Models* (John Wiley & Sons Ltd, Chichester, England, ed. 2nd, 2004), pp. 579.
34. A. Onufriev, D. Bashford, D. A. Case, *Journal of Physical Chemistry B* 104, 3712 (Apr 20, 2000).
35. R. H. Zhou, B. J. Berne, *Proceedings of the National Academy of Sciences of the United States of America* 99, 12777 (Oct 1, 2002).
36. R. Geney, M. Layten, R. Gomperts, V. Hornak, C. Simmerling, *Journal of Chemical Theory and Computation* 2, 115 (Jan-Feb, 2006).

The Nature of the Dissociation of H_2O in Water

Sarah M. Tschampel, Frank C. Pickard IV, Jennifer R. Derby, and George C. Shields*

Contribution from the Department of Chemistry, Hamilton College, 198 College Hill Road, Clinton, NY 13323

Received



Abstract:
A combined study utilizing *ab initio* theory, polarizable continuum solvation methods and thermodynamic cycles illustrates the local structure of the autoionization products of water. The present results are consistent with the most recent interpretation of X-ray absorption experiments of liquid water. The possible ionization products are limited to H_3O^+, $H_5O_2^+$, $H_9O_4^+$, $H_7O_4^-$, and OH^- for the clusters investigated ($H_3O^+(H_2O)_n$ and $OH^-(H_2O)_n$, (n=0-3)). While other ion structures may be transient structures, they are not predicted to be significant in liquid water. In addition, the formation of different ion products is shown to be dependent on the local water structure. These results are also consistent with the latest interpretations of water structure.

Introduction

Autoionization of water, or the dissociation of water into its ionic components, is one of the most fundamental reactions in chemistry. Most textbooks describe the dissociation of water one of two ways:

$$H_2O_{(l)} \rightarrow H^+_{(aq)} + OH^-_{(aq)} \qquad (1)$$

$$2H_2O_{(l)} \rightarrow H_3O^+_{(aq)} + OH^-_{(aq)} \qquad (2)$$

but the nature of the ion products is still a subject of active debate.[1-31]

The first Car-Parrinello molecular dynamics (CPMD) simulations[32] on aqueous hydroxide solutions identified $H_9O_5^-$ as the dominant structure of $OH^-_{(aq)}$.[1] The $H_9O_5^-$

$[HO^-(H_2O)_4]$ structure has four water molecules hydrogen bonded with OH^- in an approximately planar conformation.[1] Additional CPMD simulations confirmed the presence of $H_9O_5^-$ and additionally showed that at low pH, $H_3O_2^-$ is not present, while a variety of clusters can form as the pH increases.[9, 11] Spectroscopic experiments on highly basic solutions led to the prediction that $H_3O_2^-$ is the dominant solution structure of the anion.[33] The anion has also been described as a species that has three water molecules hydrogen bonded to it, $H_7O_4^-$, initially to reconcile experimental and computational studies of ester hydrolysis.[34] Subsequently, statistical mechanical quasi-chemical theory and adiabatic *ab initio* molecular dynamics simulations, contrary to CPMD results, were shown to support the prediction that $H_7O_4^-$ is the predominant structure in solution.[17, 35] A combination of neutron diffraction with hydrogen isotope substitution and Monte Carlo simulations[16] lend support to the $H_9O_5^-$ species predicted by CPMD simulations.[1] The interpretation of neutron diffraction data led to the prediction that the number of waters strongly bound to the oxygen of the hydroxyl ion is reduced from four to three as the concentration of hydroxyl ion is increased.[20, 22] A recent review on structure and dynamics of $OH^-_{(aq)}$ argues for a nonclassical, hypercoordinated solvation structure, in which $H_9O_5^-$ is the predominate ion, but $H_7O_4^-$ is also present.[26] Despite the confusion in the liquid phase, it is clear that in the gas phase three waters surround the hydroxide anion to form the first solvation shell.[14, 36] *Ab initio* calculations on gas phase $OH^-(H_2O)_n$ structures show that the $H_9O_5^-$ structure is less stable than the $H_7O_4^-$ structure with one water added to the second shell.[37, 38] The $OH^-(H_2O)$ gas phase anion is strongly bound, leading to the disappearance of the HOH intramolecular bending transition and a dramatic red-shift of the shared proton stretch.[39]

While most textbooks represent $H^+_{(aq)}$ as the hydronium ion, H_3O^+, other studies suggest $H_5O_2^+$ or $H_9O_4^+$ for the proper representation of the cation.[40-42] The dynamics of excess proton motion and proton transfer in water have been studied by both quantum and classical simulations.[1, 4-6, 24, 27, 43-48] The Zundel cation, $H_5O_2^+$, with its strongly bound proton equally shared between two waters, has been extensively studied.[15, 18, 49-52] The effects of increasing basis set size and using high levels of electron correlation for protonated systems such as $H_5O_2^+$ are well documented.[49] A combined spectroscopic/computational study predicts both the Zundel cation, $H_5O_2^+$, and the Eigen cation, $H_9O_4^+$, to be present in bulk water.[13] Protonated clusters of water molecules have been studied in the mid-IR range and the H_3O^+ and $H_5O_2^+$ ions have been observed for up to 27 water molecules.[52, 53] Studies on varying concentrations of HCl solutions using X-ray absorption combined with density functional theory have led to the conclusion that in highly acidic solutions the Eigen form dominates, while the proton is not as localized to a specific water under less acidic conditions.[25, 29] The use of multi-state empirical valence bond (MS-EVB) approaches predict that the solvation structure of the hydrated proton in water is approximately a 65:35 mixture of the Eigen and Zundel cations.[27] Ultrafast vibrational and structural dynamics of the proton in liquid water reveal that the lifetime of the $O\text{-}H^+$ stretching mode of the Eigen cation and interconversion between Eigen and Zundel hydration structures of the proton are on the order of 100 femtoseconds.[30] A neutron diffraction experiment with isotopic substitution on a concentrated HCl solution, combined with a Monte Carlo simulation, supports the existence of either hydronium ions, Eigen, or Zundel complexes.[19, 21] The accuracy of the fit between the data from Monte Carlo simulations and the diffraction data when the starting hypotheses for the

simulations are bare H^+ ions, all H_3O^+ ions, or all $H_5O_2^+$ ions in the simulation box, together with water and chlorine ions, gives the same general picture.[21] Agreement with neutron data was achieved when every H^+ ion in the simulation has at least one water molecule at a distance of 1.05 Å, which makes a hydronium ion. Three quarters of the hydronium ions have a second water at 1.32 Å, which is described as either a distorted Zundel or distorted Eigen cation depending on the other nearest water neighbors to these complexes.[21]

The nature of local water structure in the liquid is also controversial, and much experimental[54-65] and theoretical[7, 59, 66-84] research has been directed at this fundamental problem.[85-92] Benson and Siebert proposed a two-structure model for liquid water based on elegant thermodynamic arguments.[66] In their model water is a cubic octamer, which can dissociate into two planar tetramers, and they showed that isomerization between clusters reproduced the heat capacity of liquid water from 0 to 100° C.[66] Classical simulations of water dynamics predict results consistent with the force fields in use, with most predicting tetrahedral networks.[67, 75, 91] In general, water models are developed to reproduce bulk liquid properties, such as the diffusion coefficient, and the heat of vaporization. In contrast to classical treatments, quantum mechanical approaches favor a less symmetrical tetrahedral configuration of water building blocks.[73]

Recently, X-ray absorption (XAS) and X-ray Raman scattering (XRS) experiments have been used to probe the molecular arrangement in the first coordination shell of liquid water.[60, 62] In contrast to the commonly reported tetrahedral structure by both theory and experiment, the data was interpreted to be consistent with a local water structure that does not adopt a primarily tetrahedral water structure.[60, 62] Rather, each

water has two hydrogen bonded configurations, as a strong donor and as a strong acceptor. This model of water is of ring structures or linear chains embedded in a disordered water cluster network. These recent results have called into question the results from *ab initio* and classical molecular dynamics simulations.[60] Infrared spectroscopy measurements compared against the percolation model lead to the prediction that at 298 K the average coordination number is 2.4, with the dominating population consisting of molecules with three hydrogen bonds.[61]

Several subsequent studies have questioned the results from this study, and produced evidence further reaffirming the existence of tetrahedral-type networks. Raman spectra collected from 278 – 353 K combined with Monte Carlo simulations implementing a commonly used classical water model support a tetrahedral description of liquid water.[90] Recently, the Raman depolarization ratio of water as a function of temperature has been interpreted as a transition from tetrahedrally coordinated water to locally planar water structures above the phase transition temperature of 35° C.[65] In addition, two separate reinterpretations of the XAS/XRS data were shown to support tetrahedral coordination.[92, 93] Reevaluation of the H-bond criteria utilized to evaluate the XAS data brings the XAS results into agreement with other experimental data (i.e. TEY-NEXAFS).[93] Combination of the results from X-ray scattering data, which can access information outside the range of XAS, confirmed the presence of the asymmetric character observed in the XAS experiment, but not as the dominate bulk water structure. The local configuration of liquid water was still shown to be primarily that of a tetrahedrally coordinated random network.[92] In addition a combination of femtosecond 2D IR spectroscopy and molecular dynamics simulations reveal that hydrogen bonds are

only broken fleetingly in liquid water, and that nonhydrogen-bonding configurations return to a hydrogen-bonding partner within 200 fs.[89]

Recent work using first principles molecular dynamics simulations indicates that hybrid density functionals may accurately model water once longer simulation times can be reached.[80] Preliminary results suggest that each water has an average of 3.48, 3.58, and 3.67 hydrogen bonds for the PBE0, X3LYP, and B3LYP functionals.[80] The B3LYP functional performs best when compared with benchmark calculations for autoionization of a water octamer, and has been suggested as the preferable functional for Car-Parinello molecular dynamics (CPMD) and quantum mechanics/molecular mechanics simulations of autoionization of water.[94] Computation of the X-ray absorption spectrum that would result from molecular dynamics simulations leads to the prediction that the BLYP CPMD simulations do not represent the actual water structure.[82] A careful comparison of CPMD and Born-Oppenheimer molecular dynamics simulations of liquid water at a temperature of 423K and a density of 0.71 g cm^{-3} yield consistent results for the liquid structure and the self-diffusion coefficient.[95] The statistical uncertainties in the self-diffusion coefficient are quite large, and the authors point out that a CPMD simulation of 500 ps would be required to obtain a self-diffusion coefficient with a statistical uncertainty of 50%.[95] Correlation of experimental data with computed properties for liquid water is not straight forward and subject to a variety of assumptions about the local structure of bulk water and its energetic properties.[75, 83, 87] Additionally, the unique aqueous phase properties emanating from the small molecular unit of water prevail and continue to challenge both theoreticians and experimentalists alike, such as the full characterization of the autoionization of water.[59]

A few essential questions remain unanswered: for example, what exactly are the ionic states of the autoionization products of water, H^+, H_3O^+, $H_5O_2^+$, $H_7O_3^+$ or $H_9O_4^+$ and what are the corresponding anions associated with the cations? High level *ab initio* theory, polarizable continuum solvation methods and thermodyamic cycles of the dissociation of neutral water at 298.15 K were implemented to address these questions, and are reported in this paper.

Methods

The following thermodynamic cycle was utilized to evaluate the different possibilities for the dissociation of water into its corresponding cations and anions:

$$\begin{array}{ccccc} & \Delta G_{gas} & & & \\ nH_2O_{gas} & \rightarrow & (H_2O)_{n-x}OH^-_{gas} & + & (H_2O)_xH^+_{gas} \\ \downarrow \Delta G_{sol} & & \downarrow \Delta G_{sol} & & \downarrow \Delta G_{sol} \\ nH_2O_{aq} & \rightarrow & (H_2O)_{n-x}OH^-_{aq} & + & (H_2O)_xH^+_{aq} \\ & \Delta G_{aq} & & & \end{array}$$

In which ΔG_{aq} can be calculated from the thermodynamic cycle as follows:

$$\Delta G_{aq} = \Delta G_{gas} + \Delta G_{sol}[(H_2O)_{n-x}OH^-] + \Delta G_{sol}[(H_2O)_xH^+] - \Delta G_{sol}(nH_2O) \qquad (4)$$

or

$$\Delta G_{aq} = \Delta G_{gas} + \Delta\Delta G_{sol} \qquad (5)$$

where

$$\Delta\Delta G_{sol} = \Delta G_{sol}[(H_2O)_{n-x}OH^-] + \Delta G_{sol}[(H_2O)_xH^+] - \Delta G_{sol}(nH_2O) \qquad (6)$$

Since the experimental value for ΔG_{aq} is known for each possible cycle, the corresponding ion clusters can be deduced based on the fit to the experimental data. Gas phase geometry optimized structures were located at the HF/6-311++G(2d,p) and MP2/6-311++G(2df,2pd) levels as well as with the Model Chemistry Methods G2[96] and CBS-

APNO[97] using Gaussian 98.[98] All structures were verified as minima on the potential energy hypersurface through frequency calculations. Energies were computed with coupled cluster theory[99, 100] with single, double and triple substitutions.[101-105] Triple excitations were included for both the complete MP4 and CCSD(T), and all inner shell electrons were included in the electronic correlation calculations. Free energies at 298.15 K were obtained by correcting the gas phase energies with the results from the zero-point and thermal energies at the HF/6-311++G(2d,p) or MP2/6-311++G(2df,2pd) levels of theory. Additionally, gas phase values were corrected from a standard state of 1 atm to a standard state of 1 M in order to be used in the thermodynamic cycle.[106-111] The effect of using water clusters as reactants was gauged by including a water dimer, tetramer, hexamer and an octamer in addition to the usual monomer representation. Values for the free energy of solvation were calculated using the polarizable conductor model (CPCM),[112] which is based on the Polarized Continuum model (PCM).[113-119] All of the solvation computations were performed on gas phase geometries optimized at the HF/6-311++G(2d,p) level, using the 6-31G(d) and 6-31+G(d) basis sets. Additionally, the surface of each sphere was subdivided into 300 triangular tesserae, and the area of the tesserae was set to 0.3 $Å^2$.

Results

Detailed comparison of the G2, G3, CBS-QB3, and CBS-APNO model chemistry methods against experimental and high level *ab initio* predictions have verified the appropriateness of these methods for water cluster[120-122] and $H^+(H_2O)_n$ and $OH^-(H_2O)_n$ ionic cluster structure calculations.[123] Additionally, the ability of the model chemistry methods to accurately model the thermochemistry for cluster formation of neutral and

ionic clusters has been documented.[123-125] The ΔG_{sol} for $H_3O^+(H_2O)_n$ and $OH^-(H_2O)_n$ clusters, which include a maximum of three waters, and $(H_2O)_n$ (n=1,2,4,6, and 8) has been determined utilizing CPCM/HF/6-31G(d)//HF/6-311++G(2d,p) and CPCM/HF/6-31+G(d)//HF/6-311++G(2d,p), Table 1. Utilizing the values for ΔG_{sol} shown in Table 1, the $\Delta\Delta G_{sol}$ values were determined for the dissociation of water into the aforementioned ion products, Table 2. Again, the values for ΔG_{sol} shown in Table 1 were utilized to determine ΔG_{aq} for a series of reactions, in which various water clusters and monomers dissociated to form various forms of the hydronium and hydroxide ions, bottom of Table 3. The corresponding gas phase free energies, ΔG_{gas}, (using G2, CCSD(T), and CBS-APNO) have also been determined.

Table 1. ΔG_{sol} for each species investigated using structures optimized at the HF/6-311++G(2d,p) level of theory.

Specie	CPCM/HF/ 6-31G(d)	CPCM/HF/ 6-31+G(d)	Exp.[a]
H_3O^+	-105.58	-105.89	-105.00
OH^-	-112.11	-111.45	-110.00
H_2O	-6.35	-7.23	-6.31
$(H_2O)_2$	-9.56	-10.61	
$H_5O_2^+$	-76.87	-77.45	
$H_3O_2^-$	-78.21	-80.03	
$(H_2O)_4$	-7.92	-9.10	
$H_7O_3^+$	-68.68	-69.68	
$H_5O_3^-$	-74.82	-77.27	
$(H_2O)_6$	-6.08	-7.14	
$H_9O_4^+$	-56.68	-57.88	
$H_7O_4^-$	-59.39	-61.56	
$(H_2O)_8$	-4.63	-5.66	

[a]References [110, 126, 127]

Table 2. $\Delta\Delta G_{sol}$ calculated for each representation of the ionization of water.

Reaction	CPCM/HF/ 6-31G(d)	CPCM/HF/ 6-31+G(d)	Exp.
$2H_2O \rightarrow H_3O^+ + OH^-$	-204.99	-202.88	-202.38
$(H_2O)_2 \rightarrow H_3O^+ + OH^-$	-208.13	-206.73	
$4H_2O \rightarrow H_5O_2^+ + H_3O_2^-$	-129.68	-128.56	
$(H_2O)_4 \rightarrow H_5O_2^+ + H_3O_2^-$	-147.16	-148.38	
$6H_2O \rightarrow H_7O_3^+ + H_5O_3^-$	-105.40	-103.57	
$(H_2O)_6 \rightarrow H_7O_3^+ + H_5O_3^-$	-137.42	-139.81	
$8H_2O \rightarrow H_9O_4^+ + H_7O_4^-$	-65.27	-61.60	
$(H_2O)_8 \rightarrow H_9O_4^+ + H_7O_4^-$	-111.44	-113.78	
$3H_2O \rightarrow H_3O^+ + H_3O_2^-$	-164.74	-164.23	
$3H_2O \rightarrow H_5O_2^+ + OH^-$	-169.93	-167.21	
$5H_2O \rightarrow H_5O_2^+ + H_3O^+ + 2OH^-$	-374.92	-370.09	
$5H_2O \rightarrow 2H_3O^+ + OH^- + H_3O_2^-$	-369.73	-367.11	
$6H_2O \rightarrow H_3O^+ + H_5O_2^+ + OH^- + H_3O_2^-$	-334.67	-331.44	

Table 3. Free energies of ionization in the gas phase (top) and aqueous phase (bottom).

ΔG_{gas} Reaction			CCSD(T)	G2	CBS-APNO	Exp.
$2H_2O$	→	$H_3O^+ + OH^-$	227.08	225.42	225.27	226.23
$(H_2O)_2$	→	$H_3O^+ + OH^-$	227.16		225.10	
$4H_2O$	→	$H_5O_2^+ + H_3O_2^-$	177.48	177.56	177.31	
$(H_2O)_4$	→	$H_5O_2^+ + H_3O_2^-$	181.56			
$6H_2O$	→	$H_7O_3^+ + H_5O_3^-$	145.50	150.37	149.58	
$(H_2O)_6$	→	$H_7O_3^+ + H_5O_3^-$	151.24			
$8H_2O$	→	$H_9O_4^+ + H_7O_4^-$	120.49	126.16	127.60	
$(H_2O)_8$	→	$H_9O_4^+ + H_7O_4^-$	131.15			
$3H_2O$	→	$H_3O^+ + H_3O_2^-$	205.22	205.19	205.04	
$3H_2O$	→	$H_5O_2^+ + OH^-$	199.35	197.80	197.46	
$5H_2O$	→	$H_5O_2^+ + H_3O^+ + 2OH^-$	426.43	423.22	422.83	
$5H_2O$	→	$2H_3O^+ + OH^- + H_3O_2^-$	432.31	430.61	430.31	
$6H_2O$	→	$H_3O^+ + H_5O_2^+ + OH^- + H_3O_2^-$	404.56	402.98	402.58	

ΔG_{aq} Reaction			CCSD(T) [a]	CCSD(T) [b]	G2 [a]	G2 [b]	CBS-APNO [a]	CBS-APNO [b]	Exp.
$2H_2O$	→	$H_3O^+ + OH^-$	22.09	24.20	20.43	22.51	20.28	22.39	23.85
$(H_2O)_2$	→	$H_3O^+ + OH^-$	19.03	20.43			16.97	18.37	21.08
$4H_2O$	→	$H_5O_2^+ + H_3O_2^-$	47.80	48.92	47.88	49.00	47.63	48.75	28.60
$(H_2O)_4$	→	$H_5O_2^+ + H_3O_2^-$	34.40	33.18					20.70
$6H_2O$	→	$H_7O_3^+ + H_5O_3^-$	40.10	41.93	44.97	46.80	44.18	46.01	33.36
$(H_2O)_6$	→	$H_7O_3^+ + H_5O_3^-$	13.82	11.43					20.43
$8H_2O$	→	$H_9O_4^+ + H_7O_4^-$	55.22	58.89	60.89	64.56	62.33	66.00	38.12
$(H_2O)_8$	→	$H_9O_4^+ + H_7O_4^-$	19.71	17.37					20.26
$3H_2O$	→	$H_3O^+ + H_3O_2^-$	40.48	40.99	40.45	40.96	40.30	40.81	26.24
$3H_2O$	→	$H_5O_2^+ + OH^-$	29.42	32.14	27.87	30.59	27.53	30.25	26.24
$5H_2O$	→	$H_5O_2^+ + H_3O^+ + 2OH^-$	51.51	56.34	48.30	53.13	47.91	52.74	31.02
$5H_2O$	→	$2H_3O^+ + OH^- + H_3O_2^-$	62.58	65.20	60.88	63.50	60.58	63.20	31.02
$6H_2O$	→	$H_3O^+ + H_5O_2^+ + OH^- + H_3O_2^-$	69.89	73.12	68.31	71.54	62.91	71.14	33.36

All values are in kcal/mol. See text for description of determination of experimental values.
[a]CPCM/6-31G(d), [b]CPCM/6-31G+G(d)

Discussion

The experimental values for the change in free energy upon solvation of H_3O^+, OH^-, and H_2O are known,[110, 126, 127] and these values are compared with the CPCM calculated values in Table 1. Both basis sets are within reasonable agreement with the experimental value for H_3O^+, OH^-, and H_2O. The best agreement with the experimental $\Delta\Delta G_{sol}$ value is achieved from the dissociation of 2 water monomers to form a hydronium and hydroxide ion. Use of the 6-31+G(d) basis set overestimates the experimental value

by almost 2 kcal/mol for both the hydroxide ion and the water monomer. Therefore, cancellation of this error allows the CPCM/6-31+G(d) method to produce the most accurate value for $\Delta\Delta G_{sol}$ for the first reaction in Table 2. Whether this cancellation of errors is a common theme among the remaining reactions listed in Table 2 has not been verified, but since the CPCM method was parameterized using HF/6-31G(d) geometries and the value for ΔG_{sol} of water was included in the parameterization set, the error for ΔG_{sol} of water is expected to be lower for CPCM/6-31G(d). The analogous ionization from the water dimer is also in agreement with the experimental value of $\Delta\Delta G_{sol}$. Table 3 contains the values for electronic energies obtained for H_3O^+, OH^-, two water monomers, and for the water dimer with computations at the CCSD(T) level on the HF/6-311++G(2d,p) structures, and using the G2 and CBS-APNO model chemistries. This table also includes the results calculated for ΔG_{aq} obtained from combining the ΔG_{gas} quantities with the $\Delta\Delta G_{sol}$ values, shown in Table 2. The experimental value for K_c, and thus ΔG_{aq}, depends on the balanced chemical equation $2H_2O \Leftrightarrow H_3O^+ + OH^-$, yielding $K_c = [H_3O^+][OH^-]/[H_2O]^2 = K_w/[H_2O]^2$. From the value of K_w (1.0116×10^{-14} at 298 K)[128] and the concentration of water (55.34 M at 298 K) $K_c = 3.3025 \times 10^{-18}$ and ΔG_{aq} = 23.85 kcal/mol. The corresponding values of ΔG_{aq} calculated for the reaction $2H_2O \Leftrightarrow H_3O^+ + OH^-$ are 22.09, 20.43, and 20.28 kcal/mol for the CCSD(T), G2, and CBS-APNO methods, respectively, combined with the CPCM/HF/6-31G(d) solvation method. All of these numbers compare quite well to the experimental value of 23.85 kcal/mol.

Assuming that the local water structure can be represented as a more structured entity, clusters can be used in the place of monomers as the bulk water model, with the smallest possible model being the water dimer. For the reaction of the dimer, $(H_2O)_2 \Leftrightarrow$

$H_3O^+ + OH^-$, $K_c = [H_3O^+][OH^-]/[(H_2O)_2] = K_w/[(H_2O)_2]$. If bulk water is actually made up of water dimers, then the concentration of water dimers would be 27.67 M, so that $K_c = 3.65595 \times 10^{-16}$, and $\Delta G_{aq} = 21.08$ kcal/mol. Computation of ΔG_{aq} for the dissociation of the water dimer into H_3O^+ and OH^- in aqueous solution yields values for the CCSD(T) and CBS-APNO methods combined with the CPCM/6-31G(d) solvation method (and CPCM/6-31+G(d)) that deviate by 2.05 and 4.11 kcal/mol (0.65 and 2.71 kcal/mol) from the experimental value of 21.08 kcal/mol. Although these values are in worse agreement with the experimental ΔG_{aq} value than for the analogous dissociation of the water monomer, the error for these methods is still within 2.1 kcal/mol for the CCSD(T) calculations combined with either continuum solvation method. Alternative dissociations of water were explored to gain insight into both water structure and the structure of the ionic dissociation products. Utilizing various water clusters consisting of up to 8 water molecules, the only system besides the dimer that is in agreement with experiment is the dissociation of the water octamer into $H_9O_4^+$ and $H_7O_4^-$ (Table 3).

Further investigation of the dissociation of water monomers, with a maximum of 8 constituents, into various ion products reveals that one additional reaction is in agreement with experiment (taking into account the error of the computation): $3H_2O \Leftrightarrow H_5O_2^+ + OH^-$. Therefore, the $H_5O_2^+$ ion is also predicted to be formed, which agrees with experiment.[13, 21, 52, 53] The energy associated with the formation of the Zundel cation was only found to be in agreement with experiment when the local bulk water description was that of solvated monomers. Since a less structured model for liquid water is supported by recent XAS/XRS results, the description of the local water structure as solvated monomers has experimental and theoretical support.

The water octamer is an interesting model, since it does represent a primarily tetrahedral-like structure. In a D_{2d} or S_4 cubic octamer, each water in the cube forms three hydrogen bonds with neighboring waters, and each water can serve either as a hydrogen bond acceptor or donor to one additional neighboring water by virtue of the tetrahedral shape displayed by the cube.[121] Thus the formation of the $H_9O_4^+$ and $H_7O_4^-$ ions from the water octamer fits with the notion that water structure is primarily tetrahedral.

Given the errors in the methods, which lead to an uncertainty in the final free energies of several kcal/mol, these results lead us to predict that all five species, H_3O^+, $H_5O_2^+$, $H_9O_4^+$, $H_7O_4^-$, and OH^-, form from water dissociation. While other ionic species may form transiently in liquid water, the thermodynamic analysis presented here supports the view that autoionization of water is limited to the five ion products mentioned above. One caveat is that we have not looked at ion products containing more than three water molecules (i.e. $H_3O^+(H_2O)_n$ and $OH^-(H_2O)_n$ where $n > 3$). The possibility that H_3O^+ and OH^- are autoionization products, and do not always exist as tightly bound clusters, is consistent with more recent experimental studies emphasizing the dynamic nature of liquid water.[31, 60, 62]

Previous thermodynamic work on the structure of water led to a tetramer/octamer two-state model.[66] The cyclic tetramer, cyclic pentamer, and cubic octamers are very stable structures in the gas phase.[121, 122, 129, 130] The quantum chemical models employed in the present study lead to the prediction that the primary ions formed in autoionization are H_3O^+, $H_5O_2^+$, $H_9O_4^+$, $H_7O_4^-$, and OH^-. The formation of larger ions ($H_9O_4^+$ and $H_7O_4^-$) directly from water autoionization depends upon the presence of a more structured local configuration of liquid water. Of the ionization reactions investigated in this paper, only

the presence of a water octamer would result in the formation of the larger ions. The thermodynamic analysis presented here shows that formation of these two larger ions from eight solvated water monomers is not in agreement with experiment.

The use of two different models to represent the local structure of bulk water yields two different groups of ionization products. Within the representation that the local water structure can be represented as solvated monomers or dimers, H_3O^+, $H_5O_2^+$, and OH^- are the ionization products. In contrast, a loose configuration of water molecules forming a cubic octamer structure is in alignment with the idea of a tetrahedral-like local water structure, and this local structure would make the dissociation of water into the $H_9O_4^+$ and $H_7O_4^-$ ion products plausible. Based on the current interpretation of experimental data it is probable that a variety of local structures exist, with a tetrahedral-like structure dominating bulk water, yielding support for both local water descriptions described here. The view that bulk water solvates individual monomers (and possibly dimers), and is also a tetrahedral-like structure, directly limits the structures of the resulting ion products. The success of continuum solvation models for the representation of bulk water surely implies that the view of local structure as bulk water (simulated by a dielectric continuum) solvating individual monomers has a strong theoretical basis.[131] In terms of the tetrahedral-like local water structure, we note that there is a direct correlation between the increase in the strength of the local hydrogen bonding interactions, i.e. $8H_20 \rightarrow (H_2O)_8$, and the formation of larger ion clusters, i.e. $H_9O_4^+$ and $H_7O_4^-$, which also have strong local hydrogen bonding interactions. In summary, the thermodynamics that control the formation of specific ion products is dependent on the local water structure.

The ability of this method to predict the formation of the three commonly thought to exist ions, $H_5O_2^+$, $H_9O_4^+$, and $H_7O_4^-$, not only leads to additional confirmation of the presence of these ions, but also to a more detailed description of the localized water structure before and after ion formation. The structure of water predicted from these results is limited to monomers, octamers, and possibly dimers. The octamer model accounts for the mostly tetrahedral arrangement of water molecules and accounts for the formation of the $H_9O_4^+$ and $H_7O_4^-$ ion products.

The traditional Eigen structure, $H_9O_4^+$, is often described as the hydronium ion, H_3O^+, solvated by three other water molecules. Many researchers in the field actually refer to H_3O^+ as the Eigen structure,[52, 53] which can be a source of confusion. Our results suggest that the hydronium ion is constantly forming, breaking, and reforming hydrogen bonds with water molecules. Femtosecond vibrational pump-probe spectroscopy was utilized to monitor the O-H^+ stretching mode, revealing that the protonic stretching mode has a lifetime of only 120 fs.[30] These results also support a less than 100 fs transition from one hydration structure to another, which is predicted to be from the Eigen to the Zundel structure.[30] From the model implemented here, both of these structures are predicted. Additionally, the fast transition also lends support to the idea that other clusters are also likely to be present, such as the weakly coordinated hydronium ion predicted here. This result is consistent with conclusions based on CPMD simulations[4, 7, 44] and X-ray absorption experiments.[25, 29] This result aligns with the idea of proton transfer along a hydrogen bond as being essentially barrierless, because of the zero-point motion of the proton.[4, 30]

With respect to the hydroxide ion, the use of thermodynamic cycles predict two conformations, both the previously proposed configuration in which the hydroxide ion is strongly hydrogen bound to three other waters, but, additionally, that the OH^- ion itself should also be observed as more than a transient species.[17, 35] Utilization of three different density functionals in *ab initio* molecular dynamic simulations led to three models for the local configuration of the hydroxide ion: primarily $H_9O_5^-$; primarily $H_7O_4^-$; and a mixture in which $H_9O_5^-$ is the dominate configuration, but $H_7O_4^-$ also exists.[26] Only the last model, in which more than one configuration of the hydroxide ion is observed, were the relative magnitudes of the diffusion coefficients for H_3O^+, OH^-, and H_2O reproduced. Solvation of the hydroxide ion with greater than 3 explicit water molecules is outside the scope of the present study. The ability of this method to successfully distinguish the possible ionization products suggests the usefulness of this model to determine the relevance of local hydroxide configurations solvated by additional water molecules, including clusters in which the hydroxide ion behaves as a hydrogen bond donor, as suggested by the recent results from *ab initio* molecular dynamics.[26]

Conclusion

State-of-the-art quantum chemistry methods and thermodynamic cycles were implemented to investigate the autoionization of liquid water at 298 K. Based on the experimental dissociation constants for the different reactions reported in Table 3, the stable species of water and ion products were predicted from quantum computations and continuum solvation calculations. Dissociation of liquid water to form several ion products, H_3O^+, $H_5O_2^+$, $H_9O_4^+$, $H_7O_4^-$, and OH^- are found to be in agreement with current combined interpretation of the local water structure, in which there is a tetrahedral

network, but asymmetric fluctuations are ubiquitous. All other ionic species investigated here, which may exist transiently, are predicted not to exist as discrete structures.

Acknowledgments
Acknowledgment is made to the donors of The Petroleum Research Fund, administered by the ACS, to Research Corporation, to NIH, to DOD, and to Hamilton College for support of this work. This project was supported in part by NSF Grant CHE-0457275, and by NSF grants CHE-0116435 and CHE-0521063 as part of the MERCURY high-performance computer consortium (http://mercury.chem.hamilton.edu).

Supporting Information Available:
Electronic energies in Hartrees and free energy changes for all the reactions discussed in this paper. Results for CCSD(T) energy calculations performed on the MP2/6-311++G(2df,2pd) geometries, and G2 model chemistry results. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Tuckerman, M.; Laasonen, K.; Sprik, M.; Parrinello, M. *J. Chem. Phys.* 1995, 103, 150-61.
(2) Trout, B. L.; Parrinello, M. *Chem. Phys. Lett.* 1998, 288, 343-347.
(3) Trout, B. L.; Parrinello, M. *J. Phys Chem. B* 1999, 103, 7340-7345.
(4) Marx, D.; Tuckerman, M. E.; Hutter, J.; Parrinello, M. *Nature* 1999, 397, 601-604.
(5) Marx, D.; Tuckerman, M. E.; Parrinello, M. *Journal of Physics: Condensed Matter* 2000, 12, A153-A159.
(6) Day, T. J. F.; Schmitt, U. W.; Voth, G. A. *J. Am. Chem. Soc.* 2000, 122, 12027-12028.
(7) Geissler, P. L.; Dellago, C.; Chandler, D.; Hutter, J.; Parrinello, M. *Science* 2001, 291, 2121-2124.
(8) Bakker, H. J.; Nienhuys, H.-K. *Science* 2002, 297, 587-590.
(9) Chen, B.; Park, J. M.; Ivanov, I.; Tabacchi, G.; Klein, M. L.; Parrinello, M. *J. Am. Chem. Soc.* 2002, 124, 8534-8535.
(10) Tuckerman, M. E.; Marx, D.; Parrinello, M. *Nature* 2002, 417, 925-929.
(11) Chen, B.; Ivanov, I.; Park, J. M.; Parrinello, M.; Klein, M. L. *J. Phys Chem. B* 2002, 106, 12006-12016.
(12) Strajbl, M.; Hong, G.; Warshel, A. *J. Phys Chem. B* 2002, 106, 13333-13343.
(13) Kim, J.; Schmitt, U. W.; Gruetzmacher, J. A.; Voth, G. A.; Scherer, N. E. *J. Chem. Phys.* 2002, 116, 737-746.
(14) Robertson, W. H.; Diken, E. G.; Price, E. A.; Shin, J.-W.; Johnson, M. A. *Science* 2003, 299, 1367-1372.
(15) Asmis, K. R.; Pivonka, N. L.; Santambrogio, G.; Bruemmer, M.; Kaposta, C.; Neumark, D. M.; Woeste, L. *Science* 2003, 299, 1375-1377.
(16) Botti, A.; Bruni, F.; Imberti, S.; Ricci, M. A.; Soper, A. K. *J. Chem. Phys.* 2003, 119, 5001-5004.
(17) Asthagiri, D.; Pratt, L. R.; Kress, J. D.; Gomez, M. A. *Chem. Phys. Lett.* 2003, 380, 530-535.
(18) Tulub, A. A. *J. Chem. Phys.* 2004, 120, 1217-1222.
(19) Botti, A.; Bruni, F.; Imberti, S.; Ricci, M. A.; Soper, A. K. *J. Chem. Phys.* 2004, 121, 7840-7848.
(20) Botti, A.; Bruni, F.; Imberti, S.; Ricci, M. A.; Soper, A. K. *J. Chem. Phys.* 2004, 120, 10154-10162.
(21) Botti, A.; Bruni, F.; Imberti, S.; Ricci, M. A.; Soper, A. K. *J. Mol. Liq.* 2005, 117, 77-79.
(22) Botti, A.; Bruni, F.; Imberti, S.; Ricci, M. A.; Soper, A. K. *J. Mol. Liq.* 2005, 117, 81-84.
(23) Garrett, B. C., et al. *Chem. Rev.* 2005, 105, 355-389.
(24) Izvekov, S.; Voth, G. A. *J. Chem. Phys.* 2005, 123, 044505.
(25) Cavalleri, M., et al. *J. Chem. Phys.* 2006, 124, 194508.
(26) Tuckerman, M. E.; Chandra, A.; Marx, D. *Accounts of Chemical Research* 2006, 39, 151-158.
(27) Voth, G. A. *Accounts of Chemical Research* 2006, 39, 143-150.

(28) Cappa, C. D.; Smith, J. D.; Messer, B. M.; Cohen, R. C.; Saykally, R. J. *J. Phys Chem. B* 2006, 110, 5301-5309.
(29) Cappa, C. D.; Smith, J. D.; Messer, B. M.; Cohen, R. C.; Saykally, R. J. *J. Phys Chem. B* 2006, 110, 1166-1171.
(30) Woutersen, S.; Bakker, H. J. *Phys. Rev. Lett.* 2006, 96.
(31) Botti, A.; Bruni, F.; Ricci, M. A.; Soper, A. K. *J. Chem. Phys.* 2006, 125.
(32) Car, R.; Parrinello, M. *Phys. Rev. Lett.* 1985, 55, 2471-4.
(33) Librovich, N. B.; Sakun, V. P.; Sokolov, N. D. *Chem. Phys.* 1979, 39, 351-66.
(34) Zhan, C.-G.; Landry, D. W.; Ornstein, R. L. *J. Am. Chem. Soc.* 2000, 122, 2621-2627.
(35) Asthagiri, D.; Pratt, L. R.; Kress, J. D.; Gomez, M. A. *Proc. Natl. Acad. Sci.* 2004, 101, 7229-7233.
(36) Meot-Ner, M.; Speller, C. V. *J. Phys. Chem.* 1986, 90, 6616-24.
(37) Grimm, A. R.; Backskay, G. B.; Haymet, D. J. *Mol. Phys.* 1995, 86, 369-84.
(38) Novoa, J. J.; Mota, F.; del Valle, C. P.; Planas, M. *J. Phys. Chem. A* 1997, 101, 7842-7853.
(39) Roscioli, J. R.; Diken, E. G.; Johnson, M. A.; Horvath, S.; McCoy, A. B. *J. Phys. Chem. A* 2006, 110, 4943-4952.
(40) Eigen, M. *Angewandte Chemie Inter. Ed. Engl.* 1964, 3, 1-19.
(41) Zundel, G. *Hydrogen Bond* 1976, 2, 683-766.
(42) Zundel, G.; Metzger, H. *Zeitschrift fuer Physikalische Chemie (Muenchen, Germany)* 1968, 58, 225-45.
(43) Cheng, H. P.; Barnett, R. N.; Landman, U. *Chem. Phys. Lett.* 1995, 237, 161-70.
(44) Lobaugh, J.; Voth, G. A. *J. Chem. Phys.* 1996, 104, 2056-69.
(45) Ojamäe, L.; Shavitt, I.; Singer, S. J. *J. Chem. Phys.* 1998, 109, 5547-5564.
(46) Schmitt, U. W.; Voth, G. A. *Isr. J. Chem.* 1999, 39, 483-492.
(47) Schmitt, U. W.; Voth, G. A. *J. Chem. Phys.* 1999, 111, 9361-9381.
(48) Intharathep, P.; Tongraar, A.; Sagarik, K. *J. Comput. Chem.* 2006, 27, 1723-1732.
(49) Xie, Y.; Remington, R. B.; Schaefer, H. F., III *J. Chem. Phys.* 1994, 101, 4878-84.
(50) Vuilleumier, R.; Borgis, D. *Journal of Molecular Structure* 2000, 552, 117-136.
(51) Hamon, S.; Speck, T.; Mitchell, J. B. A.; Rowe, B.; Troe, J. *J. Chem. Phys.* 2005, 123.
(52) Headrick, J. M., et al. *Science* 2005, 308, 1765-1769.
(53) Shin, J. W.; Hammer, N. I.; Diken, E. G.; Johnson, M. A.; Walters, R. S.; Jaeger, T. D.; Duncan, M. A.; Christie, R. A.; Jordan, K. D. *Science* 2004, 304, 1137-1140.
(54) Woutersen, S.; Emmerichs, U.; Bakker, H. J. *Science* 1997, 278, 658-660.
(55) Keutsch, F.; Saykally, R. J. *Proc. Natl. Acad. Sci.* 2001, 98, 10533-10540.
(56) Kropman, M. F.; Bakker, H. J. *Science* 2001, 291, 2118-2120.
(57) Ruan, C.-Y.; Lobastov, V. A.; Vigliotti, F.; Chen, S.; Zewail, A. H. *Science* 2004, 304, 80-84.
(58) Myneni, S., et al. *Journal of Physics: Condensed Matter* 2002, 14, L213-L219.
(59) Bakker, H. J.; Nienhuys, H. K.; Gallot, G.; Lascoux, N.; Gale, G. M.; Leicknam, J. C.; Bratos, S. *J. Chem. Phys.* 2002, 116, 2592-2598.

(60) Wernet, P., et al. *Science* 2004, 304, 995-999.
(61) Brubach, J. B.; Mermet, A.; Filabozzi, A.; Gerschel, A.; Roy, P. *J. Chem. Phys.* 2005, 122.
(62) Näslund, L. A.; Lüning, J.; Ufuktepe, Y.; Ogasawara, H.; Wernet, P.; Bergmann, U.; Pettersson, L. G. M.; Nilsson, A. *J. Phys Chem. B* 2005, 109, 13835-13839.
(63) Öhrwall, G., et al. *J. Chem. Phys.* 2005, 123.
(64) Hakala, M.; Nygard, K.; Manninen, S.; Pettersson, L. G. M.; Hamalainen, K. *Phys. Rev. B* 2006, 73.
(65) Alphonse, N. K.; Dillon, S. R.; Dougherty, R. C.; Galligan, D. K.; Howard, L. N. *J. Phys. Chem. A* 2006, 110, 7577-7580.
(66) Benson, S. W.; Siebert, E. D. *J. Am. Chem. Soc.* 1992, 114, 4269-4276.
(67) Silvestrelli, P. L.; Parrinello, M. *J. Chem. Phys.* 1999, 111, 3572-3580.
(68) Izvekov, S.; Voth, G. A. *J. Chem. Phys.* 2002, 116, 10372-10376.
(69) Chen, B.; Ivanov, I.; Klein, M. L.; Parrinello, M. *Phys. Rev. Lett.* 2003, 91, 215503/1-215503/4.
(70) Asthagiri, D.; Pratt, L. R.; Kress, J. D. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics* 2003, 68, 041505/1-041505/7.
(71) Grossman, J. C.; Schwegler, E.; Draeger, E. W.; Gygi, F.; Galli, G. *J. Chem. Phys.* 2004, 120, 300-311.
(72) Allesch, M.; Schwegler, E.; Gygi, F.; Galli, G. *J. Chem. Phys.* 2004, 120, 5192-5198.
(73) de la Pena, L. H.; Razul, M. S. G.; Kusalik, P. G. *J. Phys. Chem. A* 2005, 109, 7236-7241.
(74) Buch, V. *J. Phys Chem. B* 2005, 109, 17771-17774.
(75) Soper, A. K. *Journal of Physics: Condensed Matter* 2005, 17, S3273-S3282.
(76) Lindenberg, A. M.; Acremann, Y.; Lowney, D. P.; Heimann, P. A.; Allison, T. K.; Matthews, T.; Falcone, R. W. *J. Chem. Phys.* 2005, 122.
(77) Prendergast, D.; Grossman, J. C.; Galli, G. *J. Chem. Phys.* 2005, 123.
(78) do Couto, P. C.; Estacio, S. G.; Cabral, B. J. C. *J. Chem. Phys.* 2005, 123.
(79) Jansen, T. L.; Hayashi, T.; Zhuang, W.; Mukamel, S. *J. Chem. Phys.* 2005, 123.
(80) Todorova, T.; Seitsonen, A. P.; Hutter, J.; Kuo, I. F. W.; Mundy, C. J. *J. Phys Chem. B* 2006, 110, 3685-3691.
(81) Fanourgakis, G. S.; Xantheas, S. S. *J. Phys. Chem. A* 2006, 110, 4100-4106.
(82) Odelius, M.; Cavalleri, M.; Nilsson, A.; Pettersson, L. G. M. *Phys. Rev. B* 2006, 73.
(83) Winter, B.; Faubel, M. *Chem. Rev.* 2006, 106, 1176-1211.
(84) Laage, D.; Hynes, J. T. *Science* 2006, 311, 832-835.
(85) Ludwig, R. *Angew. Chem. Int. Ed.* 2001, 40, 1808-1827.
(86) Bertrand, G. *J. Mol. Liq.* 2002, 101, 219-260.
(87) Head-Gordon, T.; Hura, G. *Chem. Rev.* 2002, 102, 2651-2670.
(88) Hura, G.; Russo, D.; Glaeser, R. M.; Head-Gordon, T.; Krack, M.; Parrinello, M. *PCCP Phys. Chem. Chem. Phys.* 2003, 5, 1981-1991.
(89) Eaves, J. D.; Loparo, J. J.; Fecko, C. J.; Roberts, S. T.; Tokmakoff, A.; Geissler, P. L. *Proc. Natl. Acad. Sci.* 2005, 102, 13019-13022.

(90) Smith, J. D.; Cappa, C. D.; Wilson, K. R.; Cohen, R. C.; Geissler, P. L.; Saykally, R. J. *Proc. Natl. Acad. Sci.* 2005, 102, 14171-14174.
(91) Mantz, Y. A.; Chen, B.; Martyna, G. J. *J. Phys Chem. B* 2006, 110, 3540-3554.
(92) Head-Gordon, T.; Johnson, M. E. *Proc. Natl. Acad. Sci.* 2006, 103, 7973-77.
(93) Smith, J. D.; Cappa, C. D.; Wilson, K. R.; Messer, B. M.; Cohen, R. C.; Saykally, R. J. *Science* 2004, 306, 851-853.
(94) Svozil, D.; Jungwirth, P. *J. Phys. Chem. A* 2006, 110, 9194-9199.
(95) Kuo, I. F. W.; Mundy, C. J.; McGrath, M. J.; Siepmann, J. I. *Journal of Chemical Theory and Computation* 2006, 2, 1274-1281.
(96) Curtiss, L. A.; Raghavachari, K.; Trucks, G. W.; Pople, J. A. *J. Chem. Phys.* 1991, 94, 7221-7230.
(97) Ochterski, J. W.; Petersson, G. A.; Montgomery, J. A. *J. Chem. Phys.* 1996, 104, 2598-2619.
(98) Frisch, M. J., et al. *Gaussian 98 (Revision A.11.3)*, Gaussian, Inc.: Pittsburgh, PA, 1998.
(99) Pople, J. A.; Krishnan, R.; Schlegel, H. B.; Binkley, J. S. *Int. J. Quantum Chem.* 1978, 14, 545-60.
(100) Bartlett, R. J.; Purvis, G. D. *Int. J. Quantum Chem.* 1978, 14, 561-81.
(101) Cizek, J.; Paldus, J.; Sroubkova, L. *Int. J. Quantum Chem.* 1969, 3, 149-67.
(102) Purvis, G. D., III; Bartlett, R. J. *J. Chem. Phys.* 1982, 76, 1910-18.
(103) Scuseria, G. E.; Janssen, C. L.; Schaefer, H. F., III *J. Chem. Phys.* 1988, 89, 7382-7.
(104) Scuseria, G. E.; Schaefer, H. F., III *J. Chem. Phys.* 1989, 90, 3700-3.
(105) Pople, J. A.; Head-Gordon, M.; Raghavachari, K. *J. Chem. Phys.* 1987, 87, 5968-75.
(106) Liptak, M. D.; Shields, G. C. *J. Am. Chem. Soc.* 2001, 123, 7314-7319.
(107) Liptak, M. D.; Shields, G. C. *Int. J. Quantum Chem.* 2001, 85, 727-741.
(108) Liptak, M. D.; Gross, K. C.; Seybold, P. G.; Feldgus, S.; Shields, G. C. *J. Am. Chem. Soc.* 2002, 124, 6421-6427.
(109) Palascak, M. W.; Shields, G. C. *J. Phys. Chem. A* 2004, 108, 3692-3694.
(110) Camaioni, D. M.; Schwerdtfeger, C. A. *J. Phys. Chem. A* 2005, 109, 10795-10797.
(111) Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. A* 2006, 110, 2493-2499.
(112) Barone, V.; Cossi, M. *J. Phys. Chem. A* 1998, 102, 1995-2001.
(113) Miertus, S.; Scrocco, E.; Tomasi, J. *Chem. Phys.* 1981, 55, 117-.
(114) Miertus, S.; Tomasi, J. *Chem. Phys.* 1982, 65, 239-.
(115) Cossi, M.; Barone, V.; Cammi, R.; Tomasi, J. *Chem. Phys. Lett.* 1996, 255, 327-335.
(116) Cances, E.; Mennucci, B.; Tomasi, J. *J. Chem. Phys.* 1997, 107, 3032-3041.
(117) Barone, V.; Cossi, M.; Tomasi, J. *J. Chem. Phys.* 1997, 107, 3210-3221.
(118) Barone, V.; Cossi, M.; Tomasi, J. *J. Comput. Chem.* 1998, 19, 404-417.
(119) Cossi, M.; Barone, V.; Mennucci, B.; Tomasi, J. *Chem. Phys. Lett.* 1998, 286, 253-260.
(120) Dunn, M. E.; Pokon, E. K.; Shields, G. C. *Int. J. Quantum Chem.* 2004, 100, 1065-1070.

(121) Day, M. B.; Kirschner, K. N.; Shields, G. C. *Int. J. Quantum Chem.* 2005, 102, 565-572.
(122) Day, M. B.; Kirschner, K. N.; Shields, G. C. *J. Phys. Chem. A* 2005, 109, 6773-6778.
(123) Pickard, F. C.; Pokon, E. K.; Liptak, M. D.; Shields, G. C. *J. Chem. Phys.* 2005, 122.
(124) Dunn, M. E.; Pokon, E. K.; Shields, G. C. *J. Am. Chem. Soc.* 2004, 126, 2647-2653.
(125) Pickard, F. C.; Dunn, M. E.; Shields, G. C. *J. Phys. Chem. A* 2005, 109, 4905-4910.
(126) Pliego, J. R.; Riveros, J. M. *Chem. Phys. Lett.* 2000, 332, 597-602.
(127) Pliego, J. R.; Riveros, J. M. *PCCP Phys. Chem. Chem. Phys.* 2002, 4, 1622-1627.
(128) Marshall, W. L.; Franck, E. U. *J. Phys. Chem. Ref. Data* 1981, 10, 295-304.
(129) Gruenloh, C. J.; Carney, J. R.; Arrington, C. A.; Zwier, T. S.; Fredericks, S. Y.; Jordan, K. D. *Science* 1997, 276, 1678-1681.
(130) Gruenloh, C. J.; Carney, J. R.; Hagemeister, F. C.; Arrington, C. A.; Zwier, T. S.; Fredericks, S. Y.; Wood, J. T.; Jordan, K. D. *J. Chem. Phys.* 1998, 109, 6601-6614.
(131) Cramer, C. J., *Essentials of Computational Chemistry: Theories and Models.* 2nd ed.; John Wiley & Sons, Inc.: New York, 2004; p 542.

In Search of $CS_2(H_2O)_{n=1-4}$ Clusters

Gregory M. Hartt, Timothy M. Evans, Karl N. Kirschner* and George C. Shields*

Department of Chemistry, Hamilton College, 198 College Hill Road, Clinton, NY 13323

*kkirschn@hamilton.edu; gshields@hamilton.edu

Abstract: Gaussian-3 and MP2/aug-cc-pVnZ methods have been used to calculate geometries and thermochemistry of $CS_2(H_2O)_n$, where n=1-4. An extensive molecular dynamics search followed by optimization using these two methods located two dimers, six trimers, six tetramers, and two pentamers. The MP2/aug-cc-pVDZ structure matched best with the experimental result for the $CS_2(H_2O)$ dimer, showing that diffuse functions are necessary to model the interactions found in this complex. For larger $CS_2(H_2O)_n$ clusters, the MP2/aug-cc-pVDZ minima are significantly different from the MP2(full)/6-31G* structures, revealing that the G3 model chemistry is not suitable for investigation of sulfur containing van der Waals complexes. Based on the MP2/aug-cc-pVTZ free energies, the concentration of saturated water in the atmosphere, and the average amount of CS_2 in the atmosphere, the concentrations of these clusters are predicted to be on the order of 10^4 $CS_2(H_2O)$ clusters·cm^{-3} and 10^6 $CS_2(H_2O)_2$ clusters·cm^{-3} at 298.15 K. The MP2/aug-cc-pVDZ scaled harmonic and anharmonic frequencies of the most abundant dimer cluster at 298 K are presented, along with the MP2/aug-cc-pVDZ scaled harmonic frequencies for the $CS_2(H_2O)_n$

structures predicted to be present in a low-temperature molecular beam experiment.

Introduction

A key atmospheric reaction is the oxidation of carbon disulfide by the hydroxyl radical to produce carbonyl sulfide,[1] an important atmospheric particle.[2] OCS is the most abundant sulfur molecule in the atmosphere and further oxidation of this molecule leads to the formation of acid rain.[3] Water clusters can catalyze reactions with sulfur species,[4-9] and various theoretical and experimental studies have shown the importance and existence of water clusters in the atmosphere. While thorough studies of water clusters with other important sulfur species have been performed,[18-24] we are unaware of any theoretical work on carbon disulfide and water clusters. Gaussian-3 (G3) model chemistry[25] has proven very reliable for describing water cluster structures and thermochemistry, as benchmarked against experimental and high level ab initio calculations.[17,26-31] In this paper we present the structures and energies for $CS_2(H_2O)_n$ complexes, where n=1-4, at 298.15 K. We used both G3 model chemistry and MP2 theory to investigate structure and thermochemistry, and compare them to available experimental results.[32]

Methods

Multiple structures of $CS_2(H_2O)_n$, where n=1-4, were built with SPARTAN.[33] All structures were optimized with the semi-empirical PM3 method,[34] and unique structures were further optimized using HF theory and the 6-31G* basis set.[35] Energies, enthalpies, and free energies were obtained with the G3 model chemistry. In the G3 model

chemistry, vibrational zero-point energy is obtained with a geometry optimization at the HF/6-31G* level, followed by scaling the frequencies by 0.8929. Then geometries are refined with an MP2(full)/6-31G* optimization, followed by a series of calculations used for various energetic corrections.[25]

For the $CS_2(H_2O)_n$ clusters, gas-phase molecular dynamics (MD) simulations were performed using the AMBER 8 suite of programs[36] to generate a wide range of input structures. In the first step of each MD simulation, the temperature was gradually increased from 5 K to the production temperature over one nanosecond. Production temperatures were 40 K except where n=1, where the temperature was 35 K. The production simulation lasted for nine nanoseconds. The non-bonded cutoff was 20 Å, and the GAFF force field was used.[37] The water molecules were assigned to the TIP5P water model[38] to ensure correct orientation of hydrogen bonding. The ensemble of structures from each simulation was examined, and while the simulation generated no new structures where n=1-3, the simulations produced two novel $CS_2(H_2O)_4$ configurations.

The MP2(full)/6-31G* geometries from the G3 output were used as input structures for optimizations and calculations with the MP2/aug-cc-pVDZ model.[39] Single point energy calculations were then carried out using the aug-cc-pVTZ basis set on the double zeta geometries. MP2/aug-cc-pVTZ geometry optimizations were performed on the two $CS_2(H_2O)$ and three of the $CS_2(H_2O)_2$ clusters. Anharmonic frequencies[40] of the lowest energy dimer were calculated at the MP2/aug-cc-pVDZ level. Scaled harmonic frequencies of the most likely clusters to be observed in a low-temperature molecular beam experiment were also calculated, using a scale factor of 0.9604 for the MP2/aug-cc-pVDZ minima.[41]

Results

We observed significant differences between the G3 model chemistry and MP2/aug-cc-pVDZ and MP2/aug-cc-pVTZ results. MP2/aug-cc-pVDZ and MP2/aug-cc-pVTZ calculations obtained a structure for $CS_2(H_2O)$ similar to the single experimental structure found by Ogata and Lovas,[32] while the G3 calculations gave a very different lowest-energy structure. Figure 1 exemplifies this by showing how the HF geometry optimized structure for a $CS_2(H_2O)_2$ cluster in step 1 of the G3 calculation changes slightly in step 3 of the G3 procedure, which is an MP2(full)/6-31G* optimization. More importantly, the G3 MP2(full)/6-31G* geometry, labeled C_{II}^{G3}, changes dramatically once the basis set is increased. The MP2/aug-cc-pVDZ structure that results, labeled B_{II}^{MP2}, is significantly different than the MP2(full)/6-31G* starting structure. This is a general problem with all of the minima located by these different methods; the MP2/aug-cc-pVDZ structures vary significantly from the MP2(full)/6-31G* structures, which affects the thermochemistry results. Selected optimizations at the MP2/aug-cc-pVTZ level for n=1 and n=2 structures support the aug-cc-pVDZ structures. The MP2(full)/6-31G* structures generated as part of the G3 model chemistry can be found in figures S1-S4 in the supplementary information,[42] while the MP2/aug-cc-pVDZ structures are presented in figures 2-5. We found two dimers, six trimers, six tetramers, and two pentamers. The structures are named to show order of stability, number of waters, and optimization method. The letters A–F in the figures signify stability, with A being the most stable according to the free energy at 298 K. The subscript Roman numeral signifies the number of waters in the complex, and the superscripts G3 and MP2 describes the method used to determine the structure (G3: MP2(full)/6-31G* optimization; MP2: MP2/aug-cc-pVDZ optimization). The thermochemical values for these structures can be

found in Table 1. Table 2 contains the free energies, equilibrium constants, and number of complexes per cubic centimeter predicted to be present in the atmosphere on a humid warm day. This assumes a concentration of CS_2 of 3.17 x 10^{-12} mol•L^{-1} (1.91 x 10^9 molecules•cm^{-3})[3] and a concentration of water of 1.54 x 10^{-3} mol•L^{-1} (9.30 x 10^{17} molecules•cm^{-3}),[17,43] and a temperature of 298 K. Scaled harmonic[41] and anharmonic frequencies for structure A_I^{MP2} can be found in Table 3. Tables 4-7 contain the MP2/aug-cc-pVDZ scaled harmonic frequencies for the lowest energy clusters; those that are most likely to be observable in a low-temperature molecular beam experiment. Table 8 identifies the strongest modes from these tables, modes which do not show significant overlap with water clusters. This table should aid experimentalists in identifying these clusters in their laboratories.

Table 1. ΔE_0, $\Delta H°_{298}$, and $\Delta G°_{298}$, for Incremental Association Energies using G3 and MP2 Model Chemistries at 298.15 K.[a]

	G3				MP2/aug-cc-pVDZ[b]				MP2/aug-cc-pVTZ[c]		
n	Structure	ΔE_0	$\Delta H°_{298}$	$\Delta G°_{298}$	Structure	ΔE_0	$\Delta H°_{298}$	$\Delta G°_{298}$	ΔE_0	$\Delta H°_{298}$	$\Delta G°_{298}$
	(A) $CS_2(H_2O)_{n-1} + (H_2O) \rightarrow CS_2(H_2O)_n$: Incremental Association Energetics										
1	A_I^{G3}	-0.86	-0.40	0.87	A_I^{MP2}	-0.22	0.22	4.17	0.02	0.45	4.41
	B_I^{G3}	3.83	-0.50	3.21	B_I^{MP2}	1.38	1.54	6.01	1.52	1.68	6.15
2	A_{II}^{G3}	-3.79	-3.92	3.43	A_{II}^{MP2}	0.36	1.03	3.80	0.64	1.31	4.09
	B_{II}^{G3}	-5.03	-5.31	3.90	B_{II}^{MP2}	-1.01	-1.38	5.97	-0.74	-1.11	6.24
	C_{II}^{G3}	-3.33	-3.48	4.46	C_{II}^{MP2}	-1.01	-1.58	6.80	-0.69	-1.26	7.12
	D_{II}^{G3}	-4.20	-4.48	4.68	D_{II}^{MP2}	-0.99	-1.56	6.83	-0.67	-1.24	7.15
	E_{II}^{G3}	-4.57	-4.91	4.72	E_{II}^{MP2}	0.27	-0.22	7.74	0.40	-0.09	7.87
	F_{II}^{G3}	-0.95	-0.78	6.01	F_{II}^{MP2}	0.30	-0.42	8.58	0.53	-0.19	8.81
	G_{II}^{G3}	-0.75	-0.17	6.03	G_{II}^{MP2}	2.85	3.02	8.73	2.97	3.14	8.85
3	A_{III}^{G3}	-8.96	-10.07	0.48	A_{III}^{MP2}	-2.79	-4.03	8.05	-2.48	-3.72	8.26
	B_{III}^{G3}	-7.77	-8.24	1.72	B_{III}^{MP2}	-4.78	-7.06	8.20	-4.31	-6.59	8.67
	C_{III}^{G3}	-7.51	-8.64	1.78	C_{III}^{MP2}	-3.80	-6.11	9.06	-3.49	-5.80	9.37
	D_{III}^{G3}	-7.51	-8.67	2.50	D_{III}^{MP2}	-3.48	-5.33	9.50	-3.22	-5.07	9.76
	E_{III}^{G3}	-3.61	-3.66	3.89	E_{III}^{MP2}	-2.82	-4.71	9.78	-2.51	-4.39	10.09
	F_{III}^{G3}	-0.46	0.21	4.40	F_{III}^{MP2}	-3.38	-5.74	10.67	-2.83	-5.19	11.22
4	A_{IV}^{G3}	-10.10	-10.79	-0.99	A_{IV}^{MP2}	-5.76	-7.81	5.93	-5.36	-7.41	6.33
	B_{IV}^{G3}	-9.96	-10.68	-0.40	B_{IV}^{MP2}	-5.20	-7.28	7.07	-4.80	-6.88	7.47

[a]All values are given in kilocalories per mole.
[b]G3 output served as input for MP2/aug-cc-pVDZ optimizations and frequency calculations.
[c]MP2/aug-cc-pVDZ output served as input for MP2/aug-cc-pVTZ single self-consistent field calculations.

Table 2. Free Energies[a], Equilibrium Constants, and Number of Complexes per Cubic Centimeter Predicted in the Atmosphere at 298.15 K.

n	Structure	$\Delta G°_{298}$	K_p	K_c	molec•cm^{-1}
		$CS_2(H_2O)_{n-1} + H_2O \rightarrow CS_2(H_2O)_n$			
1	A_I^{MP2}	4.41	5.89×10^{-4}	1.44×10^{-2}	4.25×10^{4}
	B_I^{MP2}	6.15	3.10×10^{-5}	7.59×10^{-4}	2.24×10^{3}
	Total				4.48×10^{4}
2	A_{II}^{MP2}	4.09	1.01×10^{-3}	6.06×10^{-1}	1.79×10^{6}
	B_{II}^{MP2}	6.24	2.67×10^{-5}	1.60×10^{-2}	4.72×10^{4}
	C_{II}^{MP2}	7.12	6.03×10^{-6}	3.61×10^{-3}	1.06×10^{4}
	D_{II}^{MP2}	7.15	5.71×10^{-6}	3.42×10^{-3}	1.01×10^{4}
	E_{II}^{MP2}	7.87	1.69×10^{-6}	1.01×10^{-3}	2.99×10^{3}
	F_{II}^{MP2}	8.81	3.49×10^{-7}	2.09×10^{-4}	6.17×10^{2}
	G_{II}^{MP2}	8.85	3.26×10^{-7}	1.95×10^{-4}	5.76×10^{2}
	Total				1.86×10^{6}

[a]Free energies (kcal/mol) calculated with MP2/aug-cc-pVTZ//MP2/aug-cc-pVDZ.

Table 3. MP2 Scaled Harmonic and Anharmonic Frequencies[a] for the $CS_2(H_2O)$ Structure A_I^{MP2}.

	MP2/aug-cc-pVDZ		Harmonic intensities		
vibrational mode[b]	scaled[c]	anharm	IR	R	Exp
ν_3(a)	3772	3736	m	s	
ν_1(a)	3645	3616	w	vs	
ν_2(a)	1559	1575	m	m	
ν_3(b)	1532	1576	vs	w	
ν_1(b)	640	648	w	vs	
in-plane ν_2(b)	364	395	w	w	
out-of-plane ν_2(b)	363	394	w	w	
rock	95	68	m	w	
wag	88	88	s	w	
intermol stretch	81	73	w	w	67[d]
out-of-plane bend	29	40	w	m	
in-plane bend	24	22	w	m	

[a]All values in cm^{-1}.
[b]Mode motion: ν_1 for symmetric stretching, ν_2 for bending, and ν_3 for asymmetric stretching. (a) signifies the water and (b) signifies the carbon disulfide.
[c]Scaling factor for MP2/aug-cc-pVDZ was 0.9604. Reference [41].
[d]Reference [32].

Table 4. MP2/aug-cc-pVDZ Scaled Harmonic Frequencies[a] for the $CS_2(H_2O)_2$ Structure B_{II}^{MP2}.

vibrational mode[b]	MP2/aug-cc-pVDZ scaled[c]	Intensity IR
ν_3(a)	3763	m
ν_3(d)	3742	m
ν_1(a)	3638	m
ν_1(d)	3571	s
ν_2(d)	1577	m
ν_2(a)	1561	m
ν_3(c)	1533	vs
ν_1(c)	641	w
ω (O-H bonded H d), twist (a)	586	m
in-plane ν_2(c)	361	w
out-of-plane ν_2(c)	361	w
symmetric ρ (d), ω (a)	338	m
asymmetric ρ (d), ω (a)	176	s
ω (S-H bonded H d), twist(a)	155	s
ρ (d), ω (a)	147	m
twist(a)	130	W
ρ (a)	93	W
intermolecular stretch (ac)	67	W
intermolecular stretch (dc)	51	W
out-of-plane intermolecular bend	20	W
in-plane intermolecular bend	17	W

[a]All values are in cm^{-1}.
[b]Mode motion: ν_1 for symmetric stretching, ν_2 for bending, ν_3 for asymmetric stretching, ω for wag, and ρ for rock. (a) signifies the water with an S···O interaction, (d) signifies the water with an OH···S interaction, and (c) signifies the carbon disulfide.
[c]MP2/aug-cc-pVDZ scaling factor was 0.9604.

Table 5. MP2/aug-cc-pVDZ Scaled Harmonic Frequencies[a] for the $CS_2(H_2O)_2$ Structure C_{II}^{MP2}.

Vibrational mode	MP2/aug-cc-pVDZ scaled[c]	Intensity IR
ν_3(d)	3754	s
ν_3(a)	3744	s
ν_1(d)	3622	m
ν_1(a)	3535	s
ν_2(a)	1578	m
ν_2(d)	1559	m
ν_3(c)	1530	vs
ν_1(c)	639	w
ω (bonded H a)	604	s
ρ (a), ω (bonded H d)	414	m
in-plane ν_2(c)	364	m
out-of-plane ν_2(c)	362	w
asymmetric ρ (d), ω (a)	209	vs
ω (free H d), twist (a)	176	m
intermolecular stretch(da)	170	m
nonconcerted ω (free H a), ρ (d)	122	s
concerted ω (free H a), ρ (d)	94	m
intermolecular stretch	80	w
intermolecular stretch	75	m
out-of-plane intermolecular bend	29	w
in-plane intermolecular bend	21	w

[a]All values are in cm^{-1}.
[b]Mode motion: ν_1 for symmetric stretching, ν_2 for bending, ν_3 for asymmetric stretching, ω for wag, and ρ for rock. (a) signifies the water with an S···O interaction, (c) signifies the carbon disulfide, and (d) signifies the water with an OH···C interaction.
[c]MP2/aug-cc-pVDZ scaling factor was 0.9604.

Table 6. MP2/aug-cc-pVDZ Scaled Harmonic Frequencies[a] for $CS_2(H_2O)_3$ Structure C_{III}^{MP2}.

vibrational mode	MP2/aug-cc-pVDZ scaled[c]	Intensity IR
ν_3(n)	3736	s
ν_3(a)	3730	m
ν_3(d)	3713	m
ν_3(bonded H d/n)	3489	vs
ν_3(bonded H d/an)	3479	vs
ν_1(bonded H dan)	3422	m
ν_2(dan)	1595	m
ν_2(dn/a)	1569	m
ν_2(d/an)	1567	s
ν_3(c)	1538	vs
ω (dan)	836	w
ν_1(c)	642	w
ω (an)	636	vs
ω (da)	551	s
ρ (da)	437	s
in-plane ν_2(c)	357	m
out-of-plane ν_2(c)	353	w
ρ (dn/a)	342	m
ρ (dan)	334	m
ω (dna)	240	m
ω (an/d)	215	m
heavy atom ring stretch	211	m
heavy atom ring distortion	193	m
heavy atom ring distortion	174	m
heavy atom ring distortion	173	w
Intermolecular stretch(ce)	68	w
Intermolecular bend(ce)	44	w
torsion(e)	39	w
twist(e)	17	w
intermolecular bend(ce)	9	w

[a]All values are in cm^{-1}

[b]Mode motion: ν_1 for symmetric stretching, ν_2 for bending, ν_3 for asymmetric stretching, ω for wag, ρ for rock, and / separates non-concerted motions. (a) signifies the water with an S···O interaction, (d) signifies the water with OH···S interaction, (n) signifies the water with no direct interaction with CS_2, (c) signifies the carbon disulfide, and (e) signifies the water trimer.

[c]MP2/aug-cc-pVDZ scaling factor was 0.9604.

Table 7. MP2/aug-cc-pVDZ Scaled Harmonic Frequencies[a] for the $CS_2(H_2O)_4$ Structure A_{IV}^{MP2}.

vibrational mode	MP2/aug-cc-pVDZ scaled[c]	Intensity IR
ν_3(a1)	3725	m
ν_3(a2)	3723	m
ν_3(free H d1/d2)	3721	m
ν_1(free H d1d2)	3721	m
ν_3(bonded H a1a2/d1d2)	3374	m
ν_3(bonded H d1/d2)	3345	vs
ν_3(bonded H a1/a2)	3320	vs
ν_1(bonded H a1a2d1d2)	3242	w
ν_2(a1a2d1d2)	1614	w
ν_2(a1a2)	1586	m
ν_2(d1d2)	1585	m
ν_2(a1a2/d1d2)	1573	s
ν_3(c)	1538	vs
ω (bonded H a1a2d1d2)	958	w
ω (bonded H a1a2)	801	m
ω (bonded H d1d2)	787	s
ω (bonded H a1a2/d1d2)	723	vs
ν_1(c)	643	w
ρ (a1a2)	442	m
ρ (d1d2)	430	w
ρ (a1a2/d1d2)	415	w
ρ (a1a2d1d2)	387	w
parallel to tetramer-plane ν_2(c)	358	w
not parallel to tetramer plane ν_2(c)	349	w
ω (free H a1a2d1d2)	284	w
ω (free H a1d1/a2)	252	s
heavy atom ring distortion	250	m
ω (free H a1d1/a2d2)	245	s
ω (a1d2/a2d1)	234	w
ω (free H d1/d2)	225	m
heavy atom ring stretch	203	w
ω (free H d1/d2)	201	m
heavy atom ring stretch	85	w
ω (free H a1a2/d1d2)	65	w
heavy atom ring distortion	61	w
intermolecular stretch(cf)	51	w
intermolecular bend(f)	46	w
intermolecular bend(cf)	23	w
intermolecular wag(cf)	9	w

[a]All values are in cm^{-1}

[b]Mode motion: ν_1 for symmetric stretching, ν_2 for bending, ν_3 for asymmetric stretching, ω for wag, ρ for rock, and / separates non-concerted motions. (a1) signifies the water with C···O interaction, (a2) signifies the water with S···O interaction, (c) signifies the carbon disulfide, (d1) signifies the water that donates to a1, (d2) signifies the water that donates to a2, and (f) signifies the water tetramer.

[c]MP2/aug-cc-pVDZ scaling factor was 0.9604.

Table 8: Intense MP2/aug-cc-pVDZ Scaled Infrared Absorptions for $CS_2(H_2O)_n$ Clusters Predicted to be Present and Observable in a Molecular Beam Low-Temperature Experimental Apparatus.[a]

Mode	A_I^{MP2}	C_{II}^{MP2}	C_{III}^{MP2}	A_{IV}^{MP2}
wag	88			
nonconcerted ω (free H a), ρ (d)		122		
ρ (da)			437	
ω (free H a1d1/a2d2)				245
ω (free H a1d1/a2)				252
ω (bonded H a1a2/d1d2)				723
ω (bonded H d1d2)				787

[a]All modes are in cm^{-1}. See previous tables for mode motion descriptions.

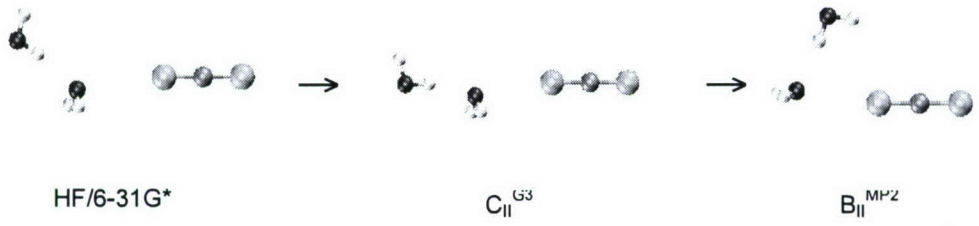


Figure 1. Structures shown left to right: HF/6-31G* optimized structure from step 1 of the G3 model chemistry, MP2(full)/6-31G* optimized structure from step 3 of the G3 model chemistry (C_{II}^{G3}), and MP2/aug-cc-pVDZ optimized structure (B_{II}^{MP2}). The HF structure was used as input for the MP2(full) optimization that produced C_{II}^{G3}, and the MP2(full) C_{II}^{G3} structure was used as input that produced the B_{II}^{MP2} structure.

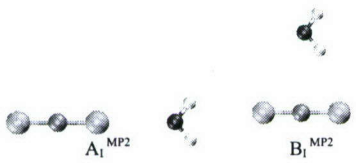


Figure 2. MP2/aug-cc-pVDZ geometries of $CS_2(H_2O)$.

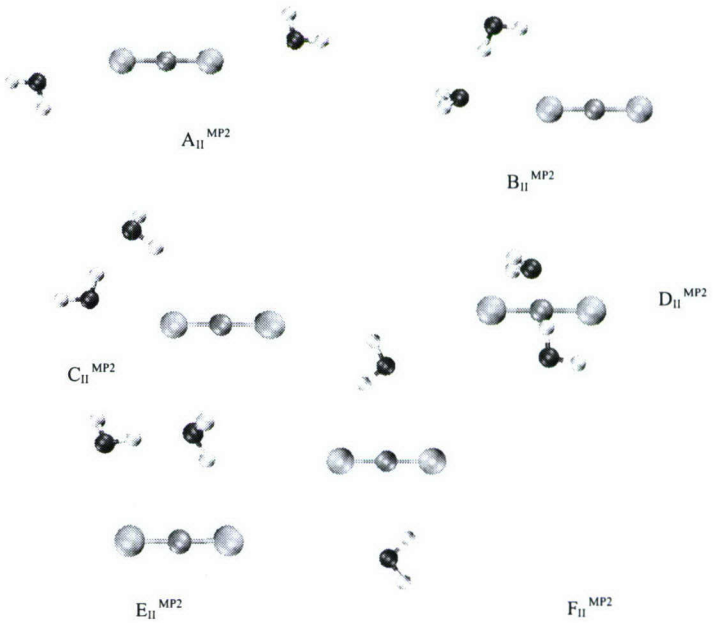
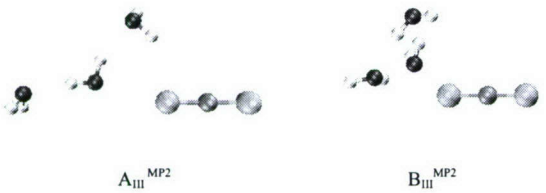


Figure 3. MP2/aug-cc-pVDZ geometries of $CS_2(H_2O)_2$.



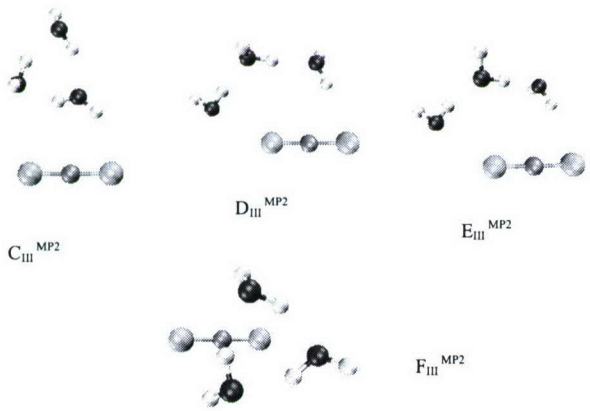


Figure 4. MP2/aug-cc-pVDZ geometries of $CS_2(H_2O)_3$.

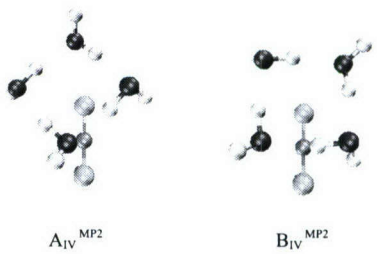


Figure 5. MP2/aug-cc-pVDZ geometries of $CS_2(H_2O)_4$.

Discussion

Structures.

Two types of structures were found for $CS_2(H_2O)$: a linear alignment of the heavy atoms found in A_I^{MP2} and B_I^{G3}, and the T-shape found in B_I^{MP2} and A_I^{G3}.[42] The linear geometry is similar to the experimental structure, obtained via Fourier transform microwave spectroscopy.[32] In the experimental structure the S···O distance is 3.197 Å, while the MP2/aug-cc-pVDZ distance is 3.116 Å and the MP2/aug-cc-pVTZ distance is 3.119 Å. Unlike the experimental structure, the MP2 structures have the hydrogen atoms in the same plane as the heavy atoms. In the experiment, the S···O-H angle is 37 degrees, which was determined assuming C_{2v} symmetry of the complex, no perturbation of the individual monomers, and that the S-C-S···O geometry is linear. Two spectroscopic constants, *B* and *C*, were determined from the microwave spectrum. The values of 1030.1109 MHz and 1026.2912 MHz are modeled fairly well by the MP2/aug-cc-pVDZ results, which are 1029.83 MHz and 1027.40 MHz. The value of *A* was assumed to be 366.4 GHz because it was the best value for the least-squares fitting procedure, although the authors comment that for a C_{2v} planar equilibrium structure that they expected the A rotational constant to be equal to 437 GHz (the *B* rotational constant for water).[32] The MP2/aug-cc-pVDZ value for *A* is 428 GHz. The experimental dipole moment is 2.078 Debye, while the calculated result is 2.802 Debye.

The linear structure using the G3 method, A_1^{G3}, is not the theoretical global energy minimum, whereas the global minimum for the MP2/aug-cc-pVDZ method is the linear structure, A_I^{MP2}. The A_1^{G3} structure, (Figure S1 in supplementary information), is qualitatively similar to the B_1^{MP2} structure displayed in Figure 2. While qualitatively the structures are similar, the two methods produce different bond lengths and angles. In the

T-shaped geometries, the S···O distances for A_I^{G3} and B_I^{MP2} are 2.972 Å and 2.754 Å, respectively. In addition to the different interaction distances, the S···H-O angle in the T-shaped structures is 125° for the MP2(full)/6-31G* geometry and 156° for the MP2/aug-cc-pVDZ structure. The addition of diffuse functions in the geometric optimizations account for the structural differences, and for the reversal of stability, by significantly altering the enthalpy and entropy of interaction.

The six structures for $CS_2(H_2O)_2$ were composed of either the low energy water dimer[26] or two separate water molecules. The G3 MP2(full)/6-31G* optimized geometries underwent significant structural changes upon the addition of diffuse functions to the basis set. Structures A_{II}^{G3} and C_{II}^{G3} (Figure S2) each contained water dimers, one perpendicular to the linear CS_2 molecule and the other extending away from one of the sulfur atoms. After MP2 optimization with diffuse functions, these waters reoriented to form the five-membered ring structures C_{II}^{MP2} and B_{II}^{MP2} (Figure 3). The global minima for both methods are similar to their respective global minima for $CS_2(H_2O)$. A_{II}^{G3} (Figure S2) consists of a four-membered ring with an extra dangling water, and A_{II}^{MP2} is a linear structure with a S···O interaction on each end similar to that of A_I^{MP2}. The diffuse functions also affected the orientation of the free (non-interacting) hydrogens in the MP2/aug-cc-pVDZ optimizations. For the A_{II}^{MP2}, B_{II}^{MP2}, C_{II}^{MP2}, and F_{II}^{MP2} structures, a free hydrogen of one or more water molecules is in the plane defined by the heavy atoms.

Optimizations of $CS_2(H_2O)_3$ resulted in two common formations of the water molecules: a linear water trimer and a cyclic water trimer, similar to the $(H_2O)_3$ global energy minimum.[26,28] In A_{III}^{G3}, the cyclic water trimer spans across CS_2 to create a S···O

van der Waals interaction with each S atom. The global minimum from MP2/aug-cc-pVDZ calculations, A_{III}^{MP2}, contains a linear water trimer bound to one sulfur atom. The C_{II}^{MP2} structure is composed of the water trimer poised above CS_2. In G3 methodology the cyclic water trimer is a key feature of these complexes' stability as three of the four lowest energy structures contain the cyclic water trimer. In the MP2/aug-cc-pVDZ calculations, however, the cyclic water trimer no longer appears in the global minimum for $CS_2(H_2O)_3$. In the structures without cyclic water trimers, the MP2/aug-cc-pVDZ optimizations placed some non-interacting hydrogens in the plane with the heavy atoms of the complex, as shown in A_{III}^{MP2}, D_{III}^{MP2}, and E_{III}^{MP2}. The G3 calculations, however, show no tendency towards a planar orientation of hydrogen in any structures. Instead, geometries such as E_{III}^{G3} and F_{III}^{G3} (Figure S3) have all non-interacting hydrogens out of the plane of the heavy atoms.

$CS_2(H_2O)_4$ shows two nearly identical structures optimized with G3 and MP2/aug-cc-pVDZ. All structures contained the cyclic S_4 water tetramer[28] centered over one of the sulfur atoms. The A_{IV}^{MP2} minimum has one of the tetramer's hydrogen bonds passing over CS_2. The C···O interaction distance in this geometry is 3.382 Å. The B_{IV}^{MP2} structure has one water of the tetramer above the CS_2, with an S···H interaction distance of 3.450 Å.

Thermochemistry.

Calculations for $CS_2(H_2O)_n$, where n=1-4, gave very different thermochemical results from the two methods, G3 and MP2/aug-cc-pVDZ. The differences are clear in Table 1 as some nearly identical structures have very different thermochemistry. The G3 results have to be discounted, as the first optimization in the G3 method uses the HF/6-

31G* method for the frequency information essential for the zero-point and entropic corrections, while the second uses the MP2(full)/6-31G* method. These two structures are often different from each other, and as shown in the previous section the lack of diffuse functions in the second geometry optimization produces quite different structures compared to MP2/aug-cc-pVDZ method. Single point energy calculations with the triple zeta basis set on the double zeta basis set geometries show a clear pattern, giving confidence that further increases of the basis set would lead to energy convergence.

The values in Table 1 show that the combination of a water molecule with a carbon disulfide molecule to form the most stable complex is only slightly endothermic; including entropy results in a 4.4 kcal·mol^{-1} free energy for formation of this complex at 298 K. Addition of another water to the A_I^{MP2} structure to form the A_{II}^{MP2} structure requires a free energy of 4.1 kcal·mol^{-1} at 298 K. Building on the A_{II}^{MP2} structure to form the A_{III}^{MP2} structure is exothermic by 3.7 kcal·mol^{-1} because of hydrogen bonding between the waters, but requires a free energy of 8.3 kcal·mol^{-1} because of the entropic cost of ordering the three water molecules on one side of the carbon disulfide molecule. Adding one more water to produce the A_{IV}^{MP2} structure is exothermic by 7.4 kcal·mol^{-1}, and requires 6.3 kcal·mol^{-1} of free energy. The S_4 water tetramer is quite stable, releasing much heat, but at the cost of greater entropy.[28]

We note that more accurate thermochemical values could be obtained by using the MP2/aug-cc-pVDZ geometries and frequencies as starting points for a G3-like calculations, or by calculating single-point energies at the CCSD(T)/aug-cc-pVnZ levels of theory. In addition, corrections for Basis Set Superposition Error would improve the thermochemical values. Such calculations would improve the absolute numbers but most

likely not change the trends, so that the MP2/aug-cc-pVTZ values effectively capture the minimum energy structures.

Based on the MP2/aug-cc-pVTZ free energies for forming the complexes in Table 1, the concentration of saturated water in the atmosphere, and the average amount of CS_2 in the atmosphere, the concentrations of the different dimer and trimer clusters are predicted in Table 2. We predict that under these conditions at 298 K, that 10^4 $CS_2(H_2O)$ clusters·cm^{-3} and 10^6 $CS_2(H_2O)_2$ clusters·cm^{-3} are present. Under the same conditions, it has been predicted that the total number of water dimers, trimers, tetramers, pentamers, and hexamers is on the order of 10^{14}, 10^{12}, 10^{11}, 10^{10}, and 10^4 clusters·cm^{-3} respectively.[17]

Frequencies.

Table 3 contains the MP2/aug-cc-pVDZ anharmonic and scaled harmonic frequencies for the A_I^{MP2} $CS_2(H_2O)$ cluster. The anharmonic frequencies are computationally expensive, and it is reassuring that there is fairly good agreement between the scaled harmonic and anharmonic frequencies in the intermolecular infrared region below 1000 cm^{-1}. Ogata and Lovas used FT microwave spectroscopy to determine the intermolecular stretch of the $CS_2(H_2O)$ cluster at 67 cm^{-1}. This frequency corresponds to the calculated values of 81 cm^{-1} (scaled MP2) and 73 cm^{-1} (anharmonic MP2) in Table 3. The strongest intensity bands, according to the MP2 scaled harmonic data, are at 1532 and 88 cm^{-1}. These two modes correspond to the asymmetric stretch of carbon disulfide (red-shifted by 10 cm^{-1}) and a wagging motion of the intermolecular complex. The 88 cm^{-1} peak is in a unique region of the spectrum relative to water clusters.[31] There is a weak intensity vibration for the water tetramer predicted at 66 cm^{-1}, and a few weak peaks below 50 cm^{-1} for the water pentamer, and all of the other intermolecular modes for

the water dimer, trimer, tetramer, and pentamer are predicted to be above 140 cm^{-1}. Thus we expect that this peak could be observed experimentally.

Based on the values for ΔE_0 in Table 1, which approximate the values for ΔG_{5K} in a cold molecular beam, we would expect the A_I, B_{II}, C_{II}, C_{III}, and A_{IV} species to predominate in a low-temperature experiment. Tables 4-8 contain the scaled harmonic frequencies for these clusters. Each vibrational mode is described in each table, along with the predicted frequency and relative IR intensity. We have previously calculated scaled harmonic and anharmonic frequencies for water clusters from the dimer up to the pentamer,[31] and we have contrasted this data with the data in tables 4-8 to determine the most likely region of the infrared spectrum to observe the $CS_2(H_2O)_n$ clusters in a laboratory experiment. Focusing on the intermolecular region, which has the least overlap with water clusters and the most reliable predicted scaled harmonic frequencies, we have listed the most likely frequencies to observe these clusters in Table 8. The C_{II}^{MP2} structure has a strong intensity mode that can best be described as a non-concerted wag of the non-interacting hydrogen atom of the water making an OH···S interaction and a rocking motion of the water making an OH···C interaction at 122 cm^{-1}. This frequency lies in between the A_I^{MP2} wag at 88 cm^{-1} and the medium intensity water dimer mode at 142 cm^{-1}, and should be detectable in a laboratory experiment. There are a few weak modes in the 128-136 cm^{-1} range for the water pentamer,[31] but at low water pressures the C_{II} frequency should be much stronger. The C_{III}^{MP2} structure has a strong intensity wagging motion of the water that makes the OH···S interaction and the water that makes the S···O interaction at 437 cm^{-1}. This vibration is in a relatively clear part of the water cluster IR spectrum, and the closest possible obfuscation of this peak would be from a predicted medium

intensity water trimer peak at 426 cm^{-1}.[31] Carbon disulfide has a weak intensity scaled MP2/aug-cc-pVDZ bending mode at 354 cm^{-1}, so the C_{III} complex's frequency is blue-shifted by approximately 80 cm^{-1}. The A_{IV}^{MP2} structure has four modes that have the potential to be observed in the laboratory. Four different wagging motions involving various parts of the intermolecular complex are predicted to be strong or very strong intensity peaks at 245, 252, 723, and 787 cm^{-1}. Carbon disulfide has a weak intensity scaled MP2/aug-cc-pVDZ symmetric stretch at 643 cm^{-1}, so the A_{IV} complex's higher frequencies are red-shifted by approximately 80 and 145 cm^{-1}. The water trimer has a weak intensity peak predicted at 229 cm^{-1}, the water pentamer has two weak intensity peaks predicted at 239 and 240 cm^{-1}, and a medium intensity peak predicted to be appear at 727 cm^{-1}.[31] Variation of water pressure and observation of the 225-275 and 700-800 cm^{-1} regions of the spectrum may allow for the detection of the A_{IV} cluster.

Methodology.

Despite the failures in the present work on van der Waals clusters of CS_2 and water, the G3 model chemistry is an excellent predictor of structures and thermochemistry for pure water clusters and ion/water clusters.[17,26-31] G3 does not use diffuse functions in its geometric optimizations, an inherent weakness for calculations of van der Waal's clusters governed by London dispersions forces. The HF/6-31G* and MP2(full)/6-31G* optimizations cannot describe the long-range interactions with the third period sulfur atom. The G3 thermochemical calculations include various correcting factors for diffuse functions, polarization functions, and other effects, which are all based on the MP2(full)/6-31G* optimized structure.[25] As the MP2(full)/6-31G* global minimum for the G3 calculations does not match the linear experimental structure for

$CS_2(H_2O)$, or the MP2/aug-cc-pVDZ structures for the $CS_2(H_2O)_n$ larger clusters, this model chemistry is not suitable for van der Waals complexes with sulfur, particularly CS_2. We note that a G3-like calculation could be performed starting with the MP2/aug-cc-pVDZ geometry and frequencies, which would most likely yield very good results.

Conclusion

We have performed an extensive molecular dynamics search for structures of $CS_2(H_2O)_n$, where n=1-4, locating two dimers, six trimers, six tetramers, and two pentamers with the G3 model chemistry and with MP2/aug-cc-pVDZ geometry optimizations. The MP2/aug-cc-pVDZ results match best with the experimental result for the $CS_2(H_2O)$ dimer, showing that diffuse functions are necessary to model the interactions found in this complex. The MP2(full)/6-31G* global minimum of $CS_2(H_2O)$ in the G3 calculations does not match the linear experimental structure for $CS_2(H_2O)$, or the MP2/aug-cc-pVDZ structure. In addition, for the larger $CS_2(H_2O)_n$ clusters, the MP2/aug-cc-pVDZ minima are significantly different from the MP2(full)/6-31G* structures. We conclude that the G3 model chemistry is not suitable for investigation of van der Waals complexes with sulfur. Based on the MP2/aug-cc-pVTZ free energies, the concentration of saturated water in the atmosphere, and the average amount of CS_2 in the atmosphere, the concentrations of $CS_2(H_2O)$ and $CS_2(H_2O)_2$ clusters are predicted to be on the order of 10^4 $CS_2(H_2O)$ clusters$\cdot$cm^{-3} and 10^6 $CS_2(H_2O)_2$ clusters$\cdot$cm^{-3} at 298 K. The most abundant clusters at 298 K are not necessarily the most abundant found in a cold molecular beam, and we predict that the A_I^{MP2}, C_{II}^{MP2}, C_{III}^{MP2}, and A_{IV}^{MP2} structures should be observable in the intermolecular infrared frequency ranges.

Acknowledgments

Acknowledgment is made to the donors of The Petroleum Research Fund, administered by the ACS, to Research Corporation, to the Camille and Henry Dreyfus Foundation, and to Hamilton College for support of this work. This project was supported in part by the U.S. Army Medical Research and Material Command's Breast Cancer Project grant W81XWH-05-1-0441, NIH grant 1R15CA115524-01, NSF grant CHE-0457275, and by NSF grants CHE-0116435 and CHE-0521063 as part of the MERCURY high-performance computer consortium (http://mercury.chem.hamilton.edu).

References

1 M. L. McKee and P. H. Wine, J. Am. Chem. Soc. **123**, 2344 (2001).

2 V. Vaida and J. E. Headrick, J. Phys. Chem. A **104** (23), 5401 (2000).

3 J. H. Seinfeld and S. N. Pandis, *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change*, 1st ed. (John Wiley & Sons, Inc., New York, 1998).

4 K. Morokuma and C. Maguruma, J. Am. Chem. Soc. **116**, 10316 (1994).

5 R. Hofmann-Sievert and A. W. Castleman, J. Phys. Chem. **88**, 3329 (1984).

6 C. E. Kolb, J. T. Jayne, D. R. Worsnop, M. L. Molina, R. F. Meads, and A. A. Viggiano, J. Am. Chem. Soc. **116**, 10314 (1994).

7 E. V. Akhmatskaya, C. J. Apps, I. H. Hillier, A. J. Masters, N. E. Watt, and J. C. Whitehead, Chem. Commun. (7), 707 (1997).

8 S. K. Ignatov, P. G. Sennikov, A. G. Razuvaev, and O. Schrems, J. Phys. Chem. A **108** (16), 3642 (2004).

9 J. M. Standard, I. S. Buckner, and D. H. Pulsifer, THEOCHEM(J. Mol. Struct.) **673** (1-3), 1 (2004).

10 S. Aloisio, J. S. Francisco, and R. R. Friedl, J. Phys. Chem. A **104** (28), 6597 (2000).

11 T. Loerting and R. L. Klaus, Proc. Natl. Acad. Sci. **97** (16), 8874 (2000).

12 T. Loerting, R. T. Kroemer, and K. R. Liedl, Chem. Commun., 999 (2000).

13 V. Vaida, J. S. Daniel, H. G. Kjaergaard, L. M. Goss, and A. F. Tuck, Q.J.R. Meteorol. Soc. **127**, 1627 (2001).

14 K. Pfeilsticker, A. Lotter, C. Peters, and H. Bösch, Science **300**, 2078 (2003).

15 A. J. Huneycutt and R. J. Saykally, Science **299**, 1329 (2003).

16 V. Vaida, H. G. Kjaergaard, and K. J. Feierabend, Int. Rev. Phys. Chem. **22** (1), 203 (2003).

17 M. E. Dunn, E. K. Pokon, and G. C. Shields, J. Am. Chem. Soc. **126** (8), 2647 (2004).

18 K. Matsumara, F. J. Lovas, and R. D. Suenram, J. Chem. Phys. **91** (10), 5887 (1989).

19 M. Planas, C. Lee, and J. J. Novoa, J. Phys. Chem. **100** (41), 16495 (1996).

20 W.-K. Li and M. L. McKee, J. Phys. Chem. A **101**, 9778 (1997).

21 E. Bishenden and D. J. Donaldson, J. Phys. Chem. A **102**, 4638 (1998).

22 M. J. Wójcik, M. Boczar, and T. A. Ford, Chem. Phys. Lett. **348**, 126 (2001).

23 Y. Tatamitani and T. Ogata, J. Chem. Phys. **121** (20), 9885 (2004).

24 A. Vila and R. A. Mosquera, Chem. Phys. **291**, 73 (2003).

[25] L. A. Curtiss, K. Raghavachari, P. C. Redfern, V. Rassolov, and J. A. Pople, J. Chem. Phys. **109** (18), 7764 (1998).
[26] M. E. Dunn, E. K. Pokon, and G. C. Shields, Int. J. Quantum Chem. **100** (6), 1065 (2004).
[27] M. B. Day, K. N. Kirschner, and G. C. Shields, Int. J. Quantum Chem. **102** (5), 565 (2005).
[28] M. B. Day, K. N. Kirschner, and G. C. Shields, J. Phys. Chem. A **109** (30), 6773 (2005).
[29] F. C. Pickard, E. K. Pokon, M. D. Liptak, and G. C. Shields, J. Chem. Phys. **122** (2) (2005).
[30] F. C. Pickard, M. E. Dunn, and G. C. Shields, J. Phys. Chem. A **109** (22), 4905 (2005).
[31] M. E. Dunn, T. M. Evans, K. N. Kirschner, and G. C. Shields, J. Phys. Chem. A **110** (1), 303 (2006).
[32] T. Ogata and F. J. Lovas, J. Mol. Spectros. **162**, 505 (1993).
[33] SPARTAN, Spartan (Wavefunction, Inc., Irvine, CA 92612, 1998).
[34] J. J. P. Stewart, J. Comput. Chem. **10** (2), 209 (1989).
[35] C. J. Cramer, *Essentials of Computational Chemistry: Theories and Models*, 2nd ed. (John Wiley & Sons, Inc., New York, 2004).
[36] D. A. Case, T. E. Cheatham, T. Darden, H. Gohlke, R. Luo, K. M. Merz, A. Onufriev, C. Simmerling, B. Wang, and R. J. Woods, J. Comput. Chem. **26** (16), 1668 (2005).
[37] J. M. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, and D. A. Case, J. Comput. Chem. **25** (9), 1157 (2004).
[38] M. W. Mahoney and W. L. Jorgensen, J. Chem. Phys. **112** (20), 8910 (2000).
[39] T. H. Dunning, J. Chem. Phys. **90** (2), 1007 (1989).
[40] V. Barone, J. Chem. Phys. **122**, 014108 (2005).
[41] P. Sinha, S. E. Boesch, C. Gu, R. A. Wheeler, and A. K. Wilson, J. Phys. Chem. A **108**, 9213 (2005).
[42] See EPAPS Document No. E-JCPS… for Tables S1-S7, which contain the type of intermolecular interaction and enthalpies for the association of CS_2 with water monomers, dimers, trimers, and tetramers for the G3 and MP2/aug-cc-pVDZ methods, and for the total association of CS_2 with two, three, or four individual water molecules. Figures S1- S4 contain the MP2(full)/6-31G* geometries obtained with the G3 method. G3, MP2/aug-cc-pVDZ, and MP2/aug-cc-pVTZ//MP2/aug-cc-pVDZ electronic energies, enthalpies, and free energies in Hartrees for all structures reported in this paper. Coordinates of all minima found with the G3, MP2/aug-cc-pVDZ, and MP2/aug-cc-pVTZ methods are presented. A direct link to this docement may be found in the online article's HTML reference section. The document may also be reached via the EPAPS homepage (http://www.aip.org/pubservs/epaps.html) or from ftp.aip.org in the directory /epaps/. See the EPAPS homepage for more information.
[43] R. B. Stull, *Meterology for Scientists and Engineers*, 2nd ed. (Brooks/Cole, Pacific Grove, CA, 2000).

Do Hydroxyl Radical−Water Clusters, $OH(H_2O)_n$, $n = 1-5$, Exist in the Atmosphere?

Marco A. Allodi, Meghan E. Dunn, Jovan Livada, Karl N. Kirschner,* and George C. Shields*

Department of Chemistry, Hamilton College, 198 College Hill Road, Clinton, New York 13323

Received: July 14, 2006; In Final Form: September 27, 2006

It has been speculated that the presence of $OH(H_2O)_n$ clusters in the troposphere could have significant effects on the solar absorption balance and the reactivity of the hydroxyl radical. We have used the G3 and G3B3 model chemistries to model the structures and predict the frequencies of hydroxyl radical/water clusters containing one to five water molecules. The reaction between hydroxyl radical clusters and methane was examined as a function of water cluster size to gain an understanding of how cluster size affects the hydroxyl radical reactivity.

Introduction

The hydroxyl radical is the most important reactive species in the troposphere. Described as the atmospheric "vacuum cleaner," the hydroxyl radical is responsible for many of the reactions that remove volatile organic compounds (VOCs) from the air.[1] The hydroxyl radical has a short lifetime in the atmosphere and is regenerated constantly. It has a global average concentration of 10^6 molecules·cm^{-3} during the daylight hours, dropping at night once photodissociation pathways that generate OH cease.[1] Because the hydroxyl radical acts as a major sink for many VOCs, its concentration influences the balance of many atmospheric species. For example, this radical oxidizes approximately 83% of annual methane emissions, supporting the fact that the hydroxyl radical is one of the most important processors of greenhouse gases. Methane concentration is increasing in the atmosphere because anthropogenic sources are more than double natural methane sources, and sources outnumber sinks by 35−40 Tg·yr^{-1}.[1]

Because it is known to form singly hydrated complexes,[2,3] the role of the $OH(H_2O)$ dimer and higher order clusters has been speculated to have an effect on atmospheric chemistry. The formation of hydrogen bonds is known to affect the spectral features and reactivity of the constituent monomers.[4] Cluster formation results in spectral peak broadening, peak shifts, and the addition of intermolecular vibrational modes that occur below 1000 cm^{-1}. The $OH(H_2O)$ dimer, recently observed by rotational spectroscopy,[2,3] is proposed as an intermediate in the interconversion of OH and H_2O,[5] and has been predicted to be a stronger oxidant than free OH.[6] The presence of $OH(H_2O)_n$ clusters could have an effect on the solar absorption balance, resulting from the presence of new absorption bands (<1000 cm^{-1}) added to the atmospheric spectrum. Additionally, the larger dipole moment of the complex enhances the intensity of absorption.[7] In a recent review article, Schrems and co-workers argue that both the hydroxyl radical's solar absorption balance and reactivity could be greatly affected by complexation with water.[4] We have undertaken this study to better understand the role that $OH(H_2O)_n$ clusters could have on atmospheric absorption and reactivity with VOCs.

* Corresponding authors. E-mail: kkirschn@hamilton.edu; gshields@hamilton.edu.

Methods

The G3 model chemistry,[8] used previously to accurately model structures and energetics for gas-phase water clusters composed of 2−8 water molecules,[9−12] ion−water clusters,[13,14] and hydroperoxy radical−water complexes,[15] has been used to model the hydroxyl radical−water complexes. Clusters were built in SPARTAN,[16] optimized using the PM3 method,[17] and followed by HF/6-31G* self-consistent field optimizations.[18] These structures served as input for G3 calculations performed using Gaussian03 versions B.02 and C.02.[19] Harmonic frequency calculations were performed on all dimer, trimer, tetramer, pentamer, and hexamer structures using the HF/6-31G* method. These frequencies were scaled by 0.8929 to obtain reliable frequency estimates for the lowest energy structures because scaled HF/6-31G* frequencies have previously been shown to be in good agreement with MP2 anharmonic and experimental values for water clusters.[20] To understand the effects of clusters on reactivity, a simple reaction between hydroxyl radical clusters and methane was examined. Reactants, transition states, and products were calculated using the G3B3 method[21] because only one geometry optimization is required with this model chemistry. In addition to the model chemistry calculations, the BHandHLYP DFT functional was used with the 6-311G** and aug-cc-pVTZ basis sets to calculate geometries and harmonic frequencies for the methane and OH radical reaction pathway. A set of single self-consistent field calculations at the CCSD(T)/aug-cc-pVTZ level were made on the BHandHLYP/aug-cc-pVTZ geometries. Kinetic rate constants were obtained using TheRate on-line program[22] using simple transition-state theory with and without Eckart tunneling.[18,23,24]

Results

The free-energy minima at 298 K for the dimer, trimer, tetramer, pentamer, and hexamer are presented in Figure 1. The relevant O−H bond lengths and hydrogen bond lengths, angles, and other geometric parameters can be found in the Supporting Information. In addition, the remaining local minima for $OH(H_2O)_n$ $n = 2-5$ clusters can be found in the Supporting Information. Table 1 contains the G3 energetics in kcal·mol^{-1} for the association of cyclic water clusters with a hydroxyl radical and for the association of hydroxyl radical−water clusters with one additional water molecule. Our reported energetics

10.1021/jp064468l CCC: $33.50
Published on Web 11/18/2006 PAGE EST: 6.3

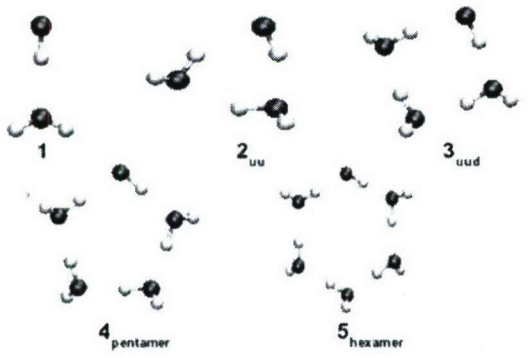


Figure 1. G3 free-energy minima at 298 K for $OH(H_2O)_n$ clusters, $n = 1-5$.

include the electronic energies of binding corrected for zero-point energy (ΔE_0), the same energies including thermal corrections ($\Delta E°_{298}$), and the enthalpies ($\Delta H°_{298}$) and Gibbs free energies ($\Delta G°_{298}$) at 298 K. Reaction i of Table 1 gives the association energetics for the formation of a hydroxyl radical−water cluster from a pure water cluster and free hydroxyl radical. Reaction ii of Table 1 gives the energetics for the sequential addition of water molecules to hydroxyl radical−water clusters.

The equilibrium constants for a standard state of 1 M (K_c), molarity (M), and the number of clusters/cm^3 predicted to be present in the lower troposphere on a humid day are presented in Table 2. These values are based on the Gibbs free energy for the association of the most stable $OH(H_2O)_{n-1}$ structure with each successive water (Table 1, reaction ii). Table 3 contains the results for the DFT calculations and for the CCSD(T) calculations for the reaction of OH and CH_4, from starting reactants to final products, including pre-reactive and post-reactive complexes, as well as the transition state for the reaction. Figure 2 illustrates the two-dimensional slice of the potential energy surface for this reaction using the data in Table 3. Table 4 contains ΔE_0, $\Delta E°_{298}$, $\Delta H°_{298}$, and $\Delta G°_{298}$ obtained using the G3B3 method for the reaction of methane with the OH radical, and the low energy C_{2v} dimer, cyclic trimer, cyclic tetramer, cyclic pentamer, and cyclic hexamer. This table also contains the G3B3 free energy of activation, $\Delta G°_{298}{}^{‡}$, for the abstraction of a hydrogen from methane by the hydroxyl radical and by $OH(H_2O)_n$ clusters and rate constants. Table 5 contains scaled frequency data for the $OH(H_2O)$ dimer and $OH(H_2O)_2$ cyclic trimer.

Discussion

In the $OH(H_2O)$ dimer global minimum, OH is the hydrogen-bond donor.[2,3,25−34] The electronic state in the G3 structure has the radical electron in the in-plane $^2A'$ configuration, which is recognized as the global minimum for $OH(H_2O)$.[2,3,28,33,34] A second electronic state is close by, with the radical electron in the out-of-plane orbital, which corresponds to the higher energy $^2A''$ state.[26,28,32−34] The excited state lies ~200 cm^{-1} above the ground state.[2] The hydrogen-bond length for the ground state of 1.912 Å compares well with the microwave values of 1.952 Å[2] and 1.945 Å.[3] The transition state for the $OH(H_2O)$- - -H- - - CH_3 structure is also in the $^2A'$ electronic state, and the transition-state structure clearly shows that the radical electron that is in-plane is forming the bond with the hydrogen that is being abstracted from methane (Supporting Information).

The low-energy structures for the $OH(H_2O)_n$ trimer, tetramer, pentamer, and hexamer clusters at 298 K are cyclic, which are similar to the analogous pure water clusters.[10,12] A previous study by Hamad, Lago, and Mejías presents structures for the dimer through hexamer using the UMP2, BLYP, and BHLYP methods.[29] They found similar structures for MP2 energy minima for the dimer, trimer, and tetramer but different energy minima for the pentamer and hexamer. Mejías and co-workers' calculations using the BHLYP method reveal similar energy minima for all cluster sizes except the pentamer.[29] Cabral do Couto and co-workers used microsolvation modeling and statistical mechanics simulations to obtain $OH(H_2O)_n$ structures where $n = 1-6$.[31] Using the MPW1PW91/aug-cc-pVDZ method, they find low-energy structures that are similar to Mejías and co-workers' findings. For the trimer through hexamer, the low-energy structure is cyclic, with the free hydrogen on the water molecules alternating in an up (u), down (d) orientation. Conversely, Mejías finds the cyclic trimer to have water molecules with an up, up (u,u) orientation and the cyclic tetramer to have an up, down, down (u,d,d) orientation. Schenter and co-workers have used CCSD(T)/aug-cc-pVTZ potential scans of the $OH(H_2O)$ potential energy surface to devise an interaction potential for the ground and excited states of the OH radical−H_2O system.[7] Their interaction potential predicts cyclic structures as the lowest energy conformers for the trimer through pentamer, and the cage structure as the most stable hexamer.[7]

By our G3 calculations, the ΔE energy minima at 0 K are the ud cyclic trimer, udu cyclic tetramer, cyclic pentamer, and cage hexamer (Table 1). This agrees with previous MPW1PW91 DFT work.[31] The same structures are found as minima after adding thermal energy to obtain $\Delta E°_{298}$, with the exception of the hexamer, for which the prism is now the lowest energy structure. Gibbs free-energy calculations at 298 K reveal the most stable configurations at that temperature. Formation of the uu cyclic trimer, uud cyclic tetramer, cyclic pentamer, and cyclic hexamer have the lowest $\Delta G°$ values at 298 K. These structures, with the exception of the cyclic pentamer, are all different from the lowest energy structures determined from $\Delta E°$. This is a common occurrence with water clusters because the zero-point corrected energies yields the right order for the free-energy surface for these clusters in a cold (5−20 K) molecular beam, while raising the temperature reorders the free-energy profile.[11,12]

Hydrogen bond lengths shorten as the cluster size increases with the longest bond length for the OH to OH_2 hydrogen bond (Supporting Information). The OH to OH_2 hydrogen bond length is longest for the dimer at 1.9 Å, shortens to 1.8 Å for the trimer, 1.73 Å for the tetramer, 1.71 Å for the pentamer, and 1.70 Å for the hexamer. The HOH to OH hydrogen bond also shortens as cluster size increases. This suggests that the OH/water electron network strengthens with additional waters. The same trend of bond lengths can be seen in pure water clusters.[10] As cluster size increases for $OH(H_2O)_n$, $n = 2-5$, the hydrogen bond angle approaches 180°. The $OH(H_2O)$ dimer has a hydrogen bond angle of 173.5°, which compresses in the more constrained trimer, then increases to nearly 180° in the cyclic pentamer and hexamer. Brauer et al. have reported that the spin dipolar constants of the dimer increase by about 33% upon complexation while the Fermi contact term changes from −73.25 MHz in free OH to −155.3 MHz in the complex.[2] Although complicated to interpret for a multielectron effect, they state unequivocally that the large change in the OH magnetic hyperfine constants upon complexation leads to a substantial change in the electron distribution of the radical upon complexation.[2] Similarly, Ohshima and co-workers have concluded from their microwave results that electron density is transferred from the water to the OH radical upon formation of the hydrogen-bonded complex.[3] Our results lead us to suggest that,

TABLE 1: G3 Energetics in kcal·mol^{-1} for the (i) Association of Cyclic Water Clusters with Hydroxyl Radical, and (ii) Incremental Association of Hydroxyl Radical−Water Clusters with an Additional Water Molecule

(i) $(H_2O)_n + OH \rightarrow HO(H_2O)_n$

(ii) $HO(H_2O)_{n-1} + H_2O \rightarrow HO(H_2O)_n$

n	ΔE_0		$\Delta E°_{298}$		$\Delta H°_{298}$		$\Delta G°_{298}$	
	i	ii	i	ii	i	ii	i	ii
1	−3.97	−3.97	−3.90	−3.90	−4.49	−4.49	1.14	1.14
2_{ud}	−6.93	−6.16	−7.58	−6.59	−8.17	−7.18	1.56	2.36
2_{uu}	−6.83	−6.05	−7.32	−6.33	−7.91	−6.92	1.37	2.17
3_{pud}	−4.84	−5.17	−4.51	−4.83	−5.10	−5.42	1.41	1.74
3_{uud}	−8.79	−9.12	−8.64	−8.97	−9.24	−9.56	−1.74	−1.38
3_{udu}	−9.03	−9.35	−9.04	−9.36	−9.63	−9.95	−1.48	−1.14
4_{S4+OH}	−5.39	−6.20	−5.32	−6.18	−5.91	−6.77	2.18	2.39
4_{D2d}	−4.87	−5.68	−4.74	−5.60	−5.33	−6.19	2.48	2.70
$4_{udu+H2O}$	−2.54	−3.36	−2.03	−2.90	−2.62	−3.49	4.10	4.32
$4_{pentamer}$	−5.39	−7.57	−5.32	−7.29	−7.02	−7.88	−0.23	−0.01
$5_{hexamer}$	−5.84	−6.32	−5.57	−5.97	−6.16	−6.56	−3.30	0.83
5_{book}	−4.56	−5.04	−4.56	−4.97	−5.16	−5.56	−0.55	3.58
$5_{pentamer+OH}$	−4.29	−4.77	−4.34	−4.75	−4.94	−5.34	−0.60	3.53
5_{cage}	−6.27	−6.75	−6.81	−7.21	−7.40	−7.80	−0.55	3.58
5_{prism}	−5.87	−6.34	−6.33	−6.73	−6.92	−7.32	−0.13	3.99

TABLE 2: Gibbs Free Energies (Standard State of 1 M) in kcal·mol^{-1}, Equilibrium Constants (Standard State of 1 M), Molarity (M), and Number of Hydroxyl Radical Clusters per Cubic Centimeter (N) Predicted to be Present in the Lower Troposphere when the Water Concentration is 0.001544 M, and the OH Concentration is 1.66058×10^{-14} M, at 298 K

reaction	$\Delta G°_{298}$	K_c	M	N
$OH + H_2O \rightarrow HO(H_2O)$	1.14	3.56	9.1×10^{-17}	5.5×10^{4}
$HO(H_2O) + H_2O \rightarrow HO(H_2O)_2$	2.17	0.631	6.4×10^{-20}	4×10^{1}
$HO(H_2O)_2 + H_2O \rightarrow HO(H_2O)_3$	−1.38	251	2.4×10^{-20}	1.4×10^{1}
$HO(H_2O)_3 + H_2O \rightarrow HO(H_2O)_4$	−0.0144	25.1	9.1×10^{-22}	5.5×10^{-1}
$HO(H_2O)_4 + H_2O \rightarrow HO(H_2O)_5$	0.828	6.05	8.5×10^{-24}	5×10^{-3}

TABLE 3: Change in Electronic Energy, Gibbs Free Energy, Enthalpy, Entropy, and Activation Energy for the Formation of CH_3 and H_2O from OH and CH_4[a]

	ΔE_0	$\Delta H°_{298}$	$\Delta G°_{298}$	$\Delta S°_{298}$	E_a[b]
		BHandHLYP/aug-cc-pVTZ[c]			
reactants → pre-reactive complex	0.20	0.72	1.43	−7.85	
reactants → transition state	8.36	7.37	13.30	−20.71	8.56
reactants → post-reactive complex	−10.04	−9.44	−7.50	−8.34	
reactants → product	−9.83	−9.42	−12.23	9.51	
		CCSD(T)/aug-cc-pVTZ[c]			
reactants → pre-reactive complex	−0.15	0.37	1.09	−2.39	
reactants → transition state	4.51	3.53	9.45	−19.87	4.72
reactants → post-reactive complex	−14.21	−13.61	−11.68	−6.50	
reactants → product	−13.46	−13.05	−15.86	9.43	
pre-reactive complex → transition state	4.66	3.16	8.36	−17.48	
transition state → post-reactive complex	−18.73	−17.14	−21.13	13.37	
post-reactive complex → product	0.75	0.56	−4.18	15.92	
		experimental data			
reactant → transition state[d]		3.89			5.08

[a] Interaction energy given in kcal·mol^{-1}, and the change in entropy given in cal·mol^{-1}·Kelvin^{-1}. ΔE_0 includes zero-point vibrational energy. [b] Activation energy determined using the equation $E_a = \Delta H°^{\ddagger}_{298} + 2RT$. [c] Single self-consistent field calculation performed on BhandHLYP/aug-cc-pVTZ optimized geometries. [d] Reference 42.

similar to pure water clusters, the cooperativity of hydrogen-bond formation distributes the partial atomic charge throughout the cluster. This effect for homodromic cyclic water clusters was first explained by Xantheas,[35] and can be seen in homodromic cyclic water structures for trimers, tetramers, pentamers, and octamers.[11,12] The cyclic structures that display cooperativity have their individual water dipoles building upon each other in the plane of the cycle, resulting in more negative charge on the oxygens and more positive charge on the hydrogens in cooperative complexes.[12] The partial atomic charges on OH, H_2O, and $OH(H_2O)$ derived from the electrostatic potential at the UMP2-(full)/6-31G* level reveal that the water hydrogens become more positive by 0.014 atomic charge in the complex, the hydroxyl hydrogen becomes more positive by 0.050 atomic charge, and the hydroxyl oxygen becomes more negative by 0.074 atomic charge.

The concentration of clusters in the atmosphere at 298 K is estimated in Table 2, for a water vapor concentration of 0.001544 M and a hydroxyl concentration of 1.66058×10^{-14} M. The equilibrium constant, K_c, (1 M standard state) is calculated from $\Delta G°$ (1 atm standard state) values from Table 1(ii). This table shows the incremental change in free energy for an existing $OH(H_2O)_n$ cluster upon incorporating another water molecule into the complex, and the abundances of specific clusters. The most abundant of the clusters, $OH(H_2O)$, is predicted to have a concentration on the order of 10^4 molecules·cm^{-3} on a warm humid day. The dimer concentration

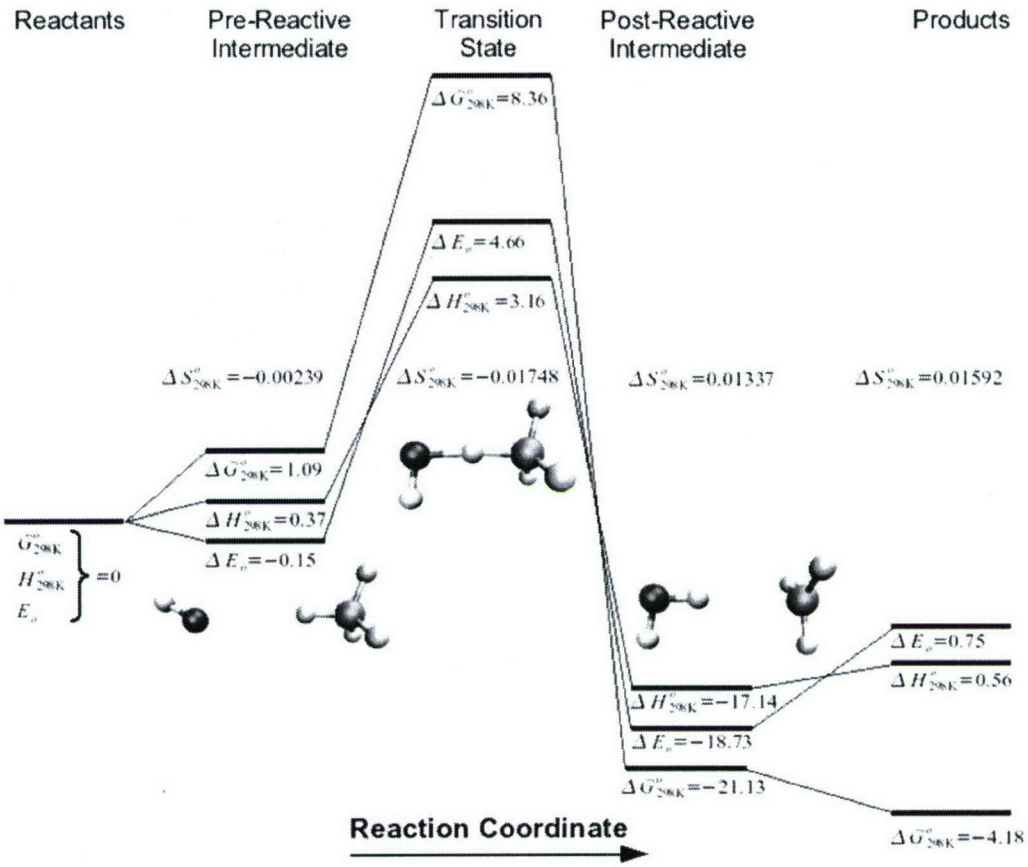


Figure 2. Two-dimensional view of the potential energy surface for the reaction of OH + CH_4 → H_2O + CH_3 using the CCSD(T)/aug-cc-pVTZ//BHandHLYP/aug-cc-pVTZ results. Energies in kcal·mol^{-1} and entropies in kcal·mol^{-1}·K^{-1}.

TABLE 4: G3B3 Change in Energy, Enthalpy, Gibbs Free Energy, and Activation Energy with Calculated Transition-State Theory Rate Constants at 298 K for the Formation of a $CH_3(H_2O)_{n+1}$ Cluster from $OH(H_2O)_n$ and CH_4[a]

	$CH_4 + OH(H_2O)_n \rightarrow CH_3{+}(H_2O)_{n+1}$						
n	ΔE_0	$\Delta E°_{298}$	$\Delta H°_{298}$	$\Delta G°_{298}$	$\Delta G°_{298}{}^{\ddagger}$	k^b	k^c
0	−13.4	−14.5	−14.0	−12.5	9.7	3.9×10^{-14}	2.0×10^{-13}
1	−13.1	−13.7	−13.7	−8.2	13.7	2.4×10^{-16}	8.7×10^{-16}
2	−14.8	−15.8	−15.4	−9.1	10.6	3.0×10^{-15}	1.2×10^{-14}
3	−14.3	−15.4	−14.9	−8.2	10.5	1.3×10^{-15}	6.9×10^{-15}
4	−14.5	−15.4	−15.1	−7.4	12.7	8.9×10^{-17}	4.1×10^{-16}
5	−14.0	−14.9	−14.6	−7.3	11.7	3.5×10^{-16}	1.5×10^{-15}

[a] Interaction energies are reported in kcal·mol^{-1}, and the rate constants are reported with units of cm^3·molecule^{-1}·s^{-1}. [b] Rate constant determined without proton tunneling. [c] Rate constant determined with Eckart proton tunneling.

is on the same order of magnitude as the water hexamer, which would be difficult to detect in the troposphere.[9] However, this does not rule out the possibility that the $OH(H_2O)_n$ clusters could have an important role in atmospheric chemistry. This is an equilibrium calculation based on an initial population for the hydroxyl radical of 10^7 radicals/cm^3. If local conditions produced a 1000th-fold increase in the OH concentration, then the OH-(H_2O) cluster concentration would increase to 10^7 clusters·cm^{-3}. Measurement of OH concentrations are complicated,[36,37] and researchers are continually trying to improve modeling and experimental methods for accurate determination of this important radical.[38−42] There is evidence for substantial variation of the concentration of hydroxyl radicals over the past several decades.[41] Field measurements and modeling studies differ between observation and predictions of OH concentration in the atmosphere, and the importance of weak electronic absorption features in contributing to significant OH production has been demonstrated.[43] On the basis of the currently accepted values for the abundances of OH, and our thermodynamic results, we do not expect clusters of $OH(H_2O)_n$ to be abundant in the atmosphere.

To put the reactivity of $OH(H_2O)_n$ clusters in context, consider Table 3, which illustrates the reactivity of the clusters by displaying the values for removing a hydrogen from methane. The oxidation of methane by the hydroxyl radical is one of the two major sources of water in the stratosphere, the other being water injection from the troposphere.[44] Experimentally, the activation barrier (E_a) for the reaction

$$CH_4 + OH \rightarrow CH_3 + H_2O$$

is 5.08 kcal·mol^{-1}, and when corrected for $2RT$ results in a $\Delta H°^{\ddagger}$ of 3.89 kcal·mol^{-1}.[45] The CCSD(T)/aug-cc-pVTZ method yields a $\Delta H°^{\ddagger}_{298}$ of 3.5 kcal·mol^{-1} and an E_a of 4.7 kcal·mol^{-1}, in agreement with the experimentally observed activation barrier. The G3B3 method is also in good agreement, producing values

TABLE 5: Vibrational Modes, Frequencies in cm^{-1}, and Relative IR Intensities for the $OH(H_2O)$ Dimer and $OH(H_2O)_2$ Cyclic Trimer from HF/6-31G* Scaled Harmonic Frequencies[a]

mode	frequency	IR intensity
	$OH(H_2O)$	
ν_3	3738	m
ν_1	3636	w
O−H stretch	3517	s
ν_2	1624	m
out-of-plane ω	541	s
in-plane ω	395	s
heavy atom stretch	173	w
τ (H_2O), ω (OH)	135	w
ω (H_2O)	122	s
	$OH(H_2O)_2$	
ν_3 (D)	3717	m
ν_3 (A)	3711	m
ν_1 (D)	3596	m
ν_1 (A)	3563	s
OH stretch	3439	s
ν_2^c (AD)	1651	m
ν_2^{nc} (A/D)	1635	m
ω^c (ADOH)	767	m
ω^{nc} (AD/OH)	550	s
ω^{nc} (A/OH)	459	s
τ^{nc} (A/D)	407	s
ρ^c (AD), ω (OH)	336	m
ρ^{nc} (A/D)	257	m
heavy atom str (AOH)	202	w
distortion	170	w
ω (A, free H)	166	m
heavy atom str (DOH)	136	w
ω (AD, free H)	100	m

[a] The relative intensity for infrared absorption (IR) are characterized as weak (w; < 40 KM/mole), medium (m; 60−160 KM/mol), or strong (s; > 195 KM/mol), and the scale factor is 0.8929. Key to mode motion: c for concerted, nc for nonconcerted, ω for wag, ρ for rock, and τ for twist. The "plane" refers to the mirror symmetry plane of the dimer. For the trimer, D and A stand for the donor and acceptor waters, respectively.

for $\Delta H°_{298}{}^{‡}$ of 3.8 kcal·mol^{-1} and for $\Delta G°_{298}{}^{‡}$ of 9.7 kcal·mol^{-1}. A two-dimensional slice of the potential energy surface for this reaction is shown in Figure 2. This curve was generated by taking the energy at each point relative to the preceding point, considering the forward reaction above and is also reported in Table 3. The curve can be described by three different types of energies: zero-point vibrationally corrected electronic energy (ΔE_0), enthalpy ($\Delta H°_{298}$), and Gibbs free energy ($\Delta G°_{298}$). A pre-reactive and post-reactive intermediate are formed on the ΔE_0 surface, whereas only the post-reactive intermediate is seen on the $\Delta H°_{298}$ surface. Formation of the pre-reactive complex from the reactants has an entropic cost, as does formation of the transition-state structure from the pre-reactive complex. Formation of the post-reactive intermediate from the transition-state structure has a favorable entropy change, as does production of the final products from the post-reactive intermediate. At room temperature and with entropy taken into account, neither intermediate is a stationary point on the $\Delta G°_{298}$ surface. The free energy of activation at room temperature is the free energy for the transition from the reactants to the transition state, 1.09 + 8.36, or 9.45 kcal·mol (Figure 2). As the temperature is lowered toward zero Kelvin, the $\Delta H°_{298}$ and $\Delta G°_{298}$ surfaces converge to the ΔE_0 surface. At very low temperatures, as found in a molecular beam, it would be possible to observe both the pre- and post-intermediates of this reaction.

Given the expense of the CCSD(T) calculations, we found that the model chemistry G3B3 method was able to provide reasonable results at a lower cost, as seen in comparing Table 4 to Table 3. As shown in the table, the hydration of OH actually reduces the Gibbs free energy released by the reaction, as well as increasing the Gibbs activation energy. In general, water clusters reduce the oxidative ability of OH with respect to methane. Thus, it appears that any water clusters that form about the hydroxyl radical will reduce, not enhance, the rate of reaction of OH with organic species in the atmosphere. We attribute this effect to the enhanced stability of $OH(H_2O)_n$ clusters as the number of waters, n, increases. As each successive water is added to make the most stable cluster, the OH to OH_2 hydrogen bond strengthens, as does the OH/water electron network. This result contradicts an earlier conclusion based on semiempirical quantum mechanical calculations.[6] Karakus and Ozkan have shown that the difference in electronic energy between transition-state structures and reactants for the reaction of $OH(H_2O)$ abstracting a hydrogen from alkanes decrease as the hydrocarbon chain is increased from one (methane) to three (propane).[46]

The experimental rate constant $CH_4 + OH \rightarrow CH_3 + H_2O$ has been determined by multiple research groups for temperature ranges from 1240 to 190 K.[47−52] The most recent experimental determination of the rate constant is $6.29 \pm 0.18 \times 10^{-15}$ to $6.8 \pm 0.14 \times 10^{-15}$ cm^3·$molecule^{-1}$·s^{-1} at 298 K.[53] A comprehensive review by Atkinson summarizes the methods and rate constants obtained from experiments from 1962 to 2002, and he recommends a value of 6.40×10^{-15} cm^3·$molecule^{-1}$·s^{-1}.[52] Using the G3B3 model chemistry results and TheRate on-line program, we determined rate constants for this reaction to be 3.9×10^{-14} cm^3·$molecule^{-1}$·s^{-1} and 2.0×10^{-13} cm^3·$molecule^{-1}$·s^{-1} when including tunneling in the rate constant calculation. Both of these values are approximately 10 and 100 times larger than the experimental value. In 1993, Melissas and Truhlar used MP2-SAC2 model chemistry results and interpolated canonical variational transition-state theory with tunneling and obtained rate constants at 298 K, $4.12−5.48 \times 10^{-15}$ cm^3·$molecule^{-1}$·s^{-1}, results that are remarkably close to today's accepted values.[23,54]

We can draw some conclusions from Table 4 if we consider that the error in the rate calculation is systematic and we only examine the relative rates as a function of hydration extent of the hydroxyl radical. Considering the tunneling-corrected rate constant, the hydration of the hydroxyl radical decreases the rate by 10- to 1000-fold, the slowest rate obtained being the hydration of the radical by a single water molecule. As discussed above, we predict that the $OH(H_2O)_n$ cluster will have an abundance of 5.5×10^4 molecules in the lower atmosphere, and a 1000-fold decrease in its reaction with methane.

The hydroxyl OH stretch of the dimer has been observed in an argon matrix at 11.5 K and assigned to bands at 3452 and 3428 cm^{-1}.[28] Later work that coupled these experiments with highly accurate anharmonic oscillator local-mode calculations revealed three bands, at 3452.2, 3428.0, and 3442.1 cm^{-1}. These three sites most likely stem from different sites within the matrix, and not from different structures, and their scaled model gives a single value of 3479.0 cm^{-1}.[33] Our gas-phase value is 3517 cm^{-1}, well within the perturbations induced by the argon matrix.[20] For spectroscopic detection of $OH(H_2O)_n$ clusters, we will focus on those vibrations that are unique and have strong intensities for the dimer and trimer because larger clusters are probably not abundant enough for detection. We have shown previously that HF/6-31G* scaled frequencies for water clusters have standard deviations of 20−25 cm^{-1} compared to experimental and to MP2 anharmonic frequencies.[20] According to the scaled HF/6-31G* frequencies, the hydroxyl radical has a vibration at 3569 cm^{-1}. The $OH(H_2O)$ dimer has an OH stretch

red shifted from the OH monomer by 52 cm^{-1}, whereas the $OH(H_2O)_2$ trimer's OH stretch is red shifted from the dimer by 78 cm^{-1}. Because of this difference in frequencies, the trimer could possibly be identified by IR spectroscopy through the 3439 cm^{-1} OH stretch. An isolated dimer could be observed at its 3517 cm^{-1} OH stretch, but this vibration would most likely be overlapped by a ν_1 $(H_2O)_3$ vibration at 3519 cm^{-1} and a ν_1 $(H_2O)_4$ vibration at 3521 cm^{-1} unless the water concentration could be reduced significantly in the IR experiment.[20]

The intermolecular vibrational modes at 541 cm^{-1} (an out-of-plane wag), 395 cm^{-1} (in-plane wag), and 122 cm^{-1} (wag) are three possible targets for detecting $OH(H_2O)$. The ν_1 vibration of $OH(H_2O)_2$ at 3563 cm^{-1}, and intermolecular vibrational modes at 550, 459, and 407 cm^{-1} have strong IR intensities that could assist in the detection of $OH(H_2O)_2$. The 550 cm^{-1} vibration is a nonconcerted wag of the acceptor and donor waters wagging out-of-phase with the OH wagging motion. This $OH(H_2O)_2$ peak could potentially be overlapped if the water dimer (vibration at 553 cm^{-1}) or water pentamer (vibration at 449 cm^{-1}) were present under experimental conditions.[20] A similar overlap with the 407 cm^{-1} vibration of $OH(H_2O)_2$, a nonconcerted twist of the acceptor water and donor water, which would be hidden by an intermolecular mode of the water trimer. If such water clusters were present in the experimental setup, then the 459 cm^{-1} vibration of $OH(H_2O)_2$ (a nonconcerted wagging of the acceptor water and hydroxyl radical) is unique from other clusters, thus a good target for detection. Another possibility for observing these complexes is to probe electronic transitions. The electronic transition $^2\Sigma^+ \leftarrow {}^2\Pi$ occurs around 308 nm, and this transition is predicted to red shift by about 2500 cm^{-1} upon formation of the $OH(H_2O)$ complex.[34]

Conclusions

The G3 model chemistry has been used to determine the structures, energetics, and frequencies of $OH(H_2O)_n$ clusters. The lowest electronic energy structures agree with previous theoretical results. Free-energy calculations at 298 K reveal that the most likely structures to be observed at room temperature usually differ from the lowest temperature structures. The lowest free-energy structures display cooperativity of hydrogen bonding, maximizing the negative partial atomic charge on oxygens and the positive partial atomic charge on hydrogens involved in hydrogen bonding in these structures. The $OH(H_2O)$ geometry is in good agreement with the experimental microwave structure. Predicted abundances of the $OH(H_2O)_n$ clusters reveal that these clusters are not predicted to be abundant in the atmosphere. Calculation of free energies of activation for the abstraction of a hydrogen from methane predict that water impedes the ability of the hydroxyl radical to oxidize organic molecules, a result that is contrary to published hypotheses.

Acknowledgment. Acknowledgment is made to the donors of the Petroleum Research Fund, administered by the ACS, to Research Corporation, and to Hamilton College for support of this work. This project was supported in part by the U.S. Army Medical Research and Material Command's Breast Cancer Project grant W81XWH-05-1-0441, NIH grant 1R15CA115524-01, NSF grant CHE-0457275, and by NSF grants CHE-0116435 and CHE-0521063 as part of the MERCURY high-performance computer consortium (http://mercury.chem.hamilton.edu).

Supporting Information Available: Electronic energies, enthalpies, and free energies in Hartrees for all structures reported in this paper. Coordinates of all minima found with the G3 and G3B3 methods. Harmonic frequencies and IR and Raman intensities of the $OH(H_2O)$ and $OH(H_2O)_2$ clusters. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Seinfeld, J. H.; Pandis, S. N. *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change*, 1st ed.; John Wiley & Sons: New York, 1998.
(2) Brauer, C. S.; Sedo, G.; Grumstrup, E. M.; Leopold, K. R.; Marshall, M. D.; Leung, H. O. *Chem. Phys. Lett.* **2005**, *401*, 420.
(3) Ohshima, Y.; Sato, K.; Sumiyoshi, Y.; Endo, Y. *J. Am. Chem. Soc.* **2005**, *127*, 1108.
(4) Sennikov, P. G.; Ignatov, S. K.; Schrems, O. *Chemphyschem* **2005**, *6*, 392.
(5) Aloisio, S.; Francisco, J. S. *Acc. Chem. Res.* **2000**, *33*, 825.
(6) Hobza, P.; Zahradnik, R. *J. Theor. Biol.* **1977**, *66*, 461.
(7) Du, S.; Francisco, J. S.; Schenter, G. K.; Iordanov, T. D.; Garrett, B. C.; Dupuis, M.; Li, J. *J. Chem. Phys.* **2006**, *124*, 224318.
(8) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Rassolov, V.; Pople, J. A. *J. Chem. Phys.* **1998**, *109*, 7764.
(9) Dunn, M. E.; Pokon, E. K.; Shields, G. C. *J. Am. Chem. Soc.* **2004**, *126*, 2647.
(10) Dunn, M. E.; Pokon, E. K.; Shields, G. C. *Int. J. Quantum Chem.* **2004**, *100*, 1065.
(11) Day, M. B.; Kirschner, K. N.; Shields, G. C. *Int. J. Quantum Chem.* **2005**, *102*, 565.
(12) Day, M. B.; Kirschner, K. N.; Shields, G. C. *J. Phys. Chem. A* **2005**, *109*, 6773.
(13) Pickard, F. C.; Pokon, E. K.; Liptak, M. D.; Shields, G. C. *J. Chem. Phys.* **2005**, *122*, 024302.
(14) Pickard, F. C.; Dunn, M. E.; Shields, G. C. *J. Phys. Chem. A* **2005**, *109*, 4905.
(15) Alongi, K. S.; Dibble, T. S.; Shields, G. C.; Kirschner, K. N. *J. Phys. Chem. A* **2006**, *110*, 3686.
(16) *SPARTAN*, 5.1.3 ed.; Wavefunction, Inc.: Irvine, CA 92612, 1998.
(17) Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 209.
(18) Cramer, C. J. *Essentials of Computational Chemistry: Theories and Models*, 2nd ed.; John Wiley & Sons: New York, 2004.
(19) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*, revision C.02; Gaussian, Inc.: Wallingford, CT, 2004.
(20) Dunn, M. E.; Evans, T. M.; Kirschner, K. N.; Shields, G. C. *J. Phys. Chem. A* **2006**, *110*, 303.
(21) Baboul, A. G.; Curtiss, L. A.; Redfern, P. C.; Raghavachari, K. *J. Chem. Phys.* **1999**, *110*, 7650.
(22) Truong, T. N.; Nayak, M.; Huynh, H. H.; Cook, T.; Mahajan, P.; Tran, L. T.; Bharath, J.; Jain, S.; Pham, H. B.; Boonyasiriwat, C.; Nguyen, N.; Andersen, E.; Kim, Y.; Choe, S.; Choi, J.; Cheatham, T. E.; Facelli, J. C. *J. Chemical Inf. Model.* **2006**, *46*, 971.
(23) Truong, T. N.; Truhlar, D. G. *J. Chem. Phys.* **1990**, *93*, 1761.
(24) Truhlar, D. G.; Garrett, B. C.; Klippenstein, S. J. *J. Phys. Chem.* **1996**, *100*, 12771.
(25) Nanayakkara, A. A.; Balintkurti, G. G.; Williams, I. H. *J. Phys. Chem.* **1992**, *96*, 3662.
(26) Xie, Y. M.; Schaefer, H. F. *J. Chem. Phys.* **1993**, *98*, 8829.
(27) Wang, B. S.; Hou, H.; Gu, Y. S. *Chem. Phys. Lett.* **1999**, *303*, 96.
(28) Langford, V. S.; McKinley, A. J.; Quickenden, T. I. *J. Am. Chem. Soc.* **2000**, *122*, 12859.
(29) Hamad, S.; Lago, S.; Mejias, J. A. *J. Phys. Chem. A* **2002**, *106*, 9104.
(30) Zhou, Z. Y.; Qu, Y. H.; Fu, A. P.; Du, B. N.; He, F. X.; Gao, H. W. *Int. J. Quantum Chem.* **2002**, *89*, 550.
(31) do Couto, P. C.; Guedes, R. C.; Cabral, B. J. C.; Simoes, J. A. M. *J. Chem. Phys.* **2003**, *119*, 7344.

(32) Engdahl, A.; Karlstrom, G.; Nelander, B. *J. Chem. Phys.* **2003**, *118*, 7797.
(33) Cooper, P. D.; Kjaergaard, H. G.; Langford, V. S.; McKinley, A. J.; Quickenden, T. I.; Schofield, D. P. *J. Am. Chem. Soc.* **2003**, *125*, 6048.
(34) Schofield, D. P.; Kjaergaard, H. G. *J. Chem. Phys.* **2004**, *120*, 6930.
(35) Xantheas, S. S. *NATO ASI Ser., Ser. C* **2000**, *561*, 119.
(36) Heard, D. E.; Pilling, M. J. *Chem. Rev.* **2003**, *103*, 5163.
(37) Heard, D. E. *Annu. Rev. Phys. Chem.* **2006**, *57*, 191.
(38) Smith, S. C.; Lee, J. D.; Bloss, W. J.; Johnson, G. P.; Ingham, T.; Heard, D. E. *Atmos. Chem. Phys.* **2006**, *6*, 1435.
(39) Krol, M.; van Leeuwen, P. J.; Lelieveld, J. *J. Geophys. Res., [Atmos.]* **1998**, *103*, 10697.
(40) Krol, M.; van Leeuwen, P. J.; Lelieveld, J. *J. Geophys. Res., [Atmos.]* **2001**, *106*, 23159.
(41) Prinn, R. G.; Huang, J.; Weiss, R. F.; Cunnold, D. M.; Fraser, P. J.; Simmonds, P. G.; McCulloch, A.; Harth, C.; Salameh, P.; O'Doherty, S.; Wang, R. H. J.; Porter, L.; Miller, B. R. *Science* **2001**, *292*, 1882.
(42) Prinn, R. G.; Huang, J. *J. Geophys. Res., [Atmos.]* **2001**, *106*, 23151.
(43) Matthews, J.; Sinha, A.; Francisco, J. S. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 7449.
(44) Remsberg, E. E.; Bhatt, P. P.; Russell, J. M. *J. Geophys. Res., [Atmos.]* **1996**, *101*, 6749.
(45) DeMore, W. B.; Sander, S. P.; Howard, C. J.; Ravishankara, A. R.; Goldern, D. M.; Kolb, C. E.; Hampson, R. F.; Kurylo, M. J.; Molina, M. J. *Chemical Kinetics and Photochemical Data for Use in Stratospheric Modeling*; NASA/Jet Propulsion Laboratory, 1997.
(46) Karakus, N.; Ozkan, R. *J. Mol. Struct.: THEOCHEM* **2005**, *724*, 39.
(47) Vaghjiani, G. L.; Ravishankara, A. R. *Nature (London)* **1991**, *350*, 406.
(48) Finlaysonpitts, B. J.; Ezell, M. J.; Jayaweera, T. M.; Berko, H. N.; Lai, C. C. *Geophys. Res. Lett.* **1992**, *19*, 1371.
(49) Saunders, S. M.; K. J., H.; Pilling, K. J.; Baulch, M. J.; Smurthwaite, P. I. *Optical Methods in Atmospheric Chemistry*; SPIE: Berlin, 1992; Vol. 1715.
(50) Dunlop, J. R.; Tully, F. P. *J. Phys. Chem.* **1993**, *97*, 11148.
(51) Gierczak, T.; Talukdar, R. K.; Herndon, S. C.; Vaghjiani, G. L.; Ravishankara, A. R. *J. Phys. Chem. A* **1997**, *101*, 3125.
(52) Atkinson, R. *Atmos. Chem. Phys.* **2003**, *3*, 2233.
(53) Bryukov, M. G.; Knyazev, V. D.; Lomnicki, S. M.; McFerrin, C. A.; Dellinger, B. *J. Phys. Chem. A* **2004**, *108*, 10464.
(54) Melissas, V. S.; Truhlar, D. G. *J. Chem. Phys.* **1993**, *99*, 1013.

CCSD(T), W1, and Other Model Chemistry Predictions for Gas-Phase Deprotonation Reactions

FRANK C. PICKARD IV, DANIEL R. GRIFFITH, SKYLAR J. FERRARA, MATTHEW D. LIPTAK, KARL N. KIRSCHNER, GEORGE C. SHIELDS

Department of Chemistry, Hamilton College, 198 College Hill Road, Clinton, New York 13323

Received 27 February 2006; accepted 25 April 2006
Published online 24 July 2006 in Wiley InterScience (www.interscience.wiley.com).
DOI 10.1002/qua.21105

ABSTRACT: A series of CCSD(T) single-point calculations on MP4(SDQ) geometries and the W1 model chemistry method have been used to calculate $\Delta H°$ and $\Delta G°$ values for the deprotonation of 17 gas-phase reactions where the experimental values have reported accuracies within 1 kcal/mol. These values have been compared with previous calculations using the G3 and CBS model chemistries and two DFT methods. The most accurate CCSD(T) method uses the aug-cc-pVQZ basis set. Extrapolation of the aug-cc-pVTZ and aug-cc-pVQZ results yields the most accurate agreement with experiment, with a standard deviation of 0.58 kcal/mol for $\Delta G°$ and 0.70 kcal/mol for $\Delta H°$. Standard deviations from experiment for $\Delta G°$ and $\Delta H°$ for the W1 method are 0.95 and 0.83 kcal/mol, respectively. The G3 and CBS-APNO results are competitive with W1 and are much less expensive. Any of the model chemistry methods or the CCSD(T)/aug-cc-pVQZ method can serve as a valuable check on the accuracy of experimental data reported in the National Institutes of Standards and Technology (NIST) database.
 Int J Quantum Chem 106: 3122–3128, 2006

Introduction

The ability of modern quantum chemical methods to calculate gas-phase reaction and free energies accurately is essential for a variety of chemical predictions, such as the accurate a priori calculation of gas-phase equilibrium constants. Computed $\Delta G°$ values must be correct to within 1 kcal/mol to achieve chemical accuracy. We have recently evaluated the ability of the Gaussian (G*n*) model chemistries [1, 2] and the complete basis set (CBS) model chemistries [3–5] to model gas-phase deprotonation accurately, by examining their performance against 17 reactions in the National Insti-

Correspondence to: G. Shields; e-mail: gshields@hamilton.edu
Contract grant sponsor: DOD.
Contract grant number: W81XWH-05-1-0441.
Contract grant sponsor: NIH.
Contract grant number: 1R15CA115524-01.
Contract grant sponsor: Hamilton College.
Contract grant sponsor: NSF.
Contract grant number: CHE-0457275.
Contract grant sponsor: NSF.
Contract grant numbers: CHE-0116435; CHE-0521063.
Contract grant sponsor: ACS PRF.
Contract grant number: 37339-B6.
Contract grant sponsor: Research Corporation.
Contract grant number: CC6041.

tutes of Standards and Technology (NIST) database [6], where the deprotonation $\Delta G°$ is reported as highly accurate (defined as an experimental error of <1 kcal/mol) [7]. The high computational cost of the model chemistries led us to evaluate a large range of density functional theory methods against the same experimental database [8]. Additionally, a search for the most accurate methods available, combined with a desire to check the experimental data, led us to determine energies with the CCSD(T)/aug-cc-pV*n*Z model using MP4(SDQ) geometries. In the present work, we report on the performance of CCSD(T) [9] and the W1 model chemistries [10] against this experimental data set [6].

Methods

We used the W1 [10] model chemistry and CCSD(T) [9] (coupled cluster with all single and double substitutions along with a quasi-perturbative treatment of connected triple excitations) implemented within Gaussian 03 [11] and Gaussian 98 [12]. The coupled cluster calculations included triple excitations for both the complete fourth-order Møller–Plesset (MP4) and the CCSD(T) energies (E4T keyword). The coupled cluster single-point energy calculations used the augmented correlation consistent polarized *n*-tuple zeta basis sets (aug-cc-pV*n*Z, n = D, T, Q, 5) of Dunning and coworkers [13–15]. These calculations were performed on geometries obtained using MP4 perturbation theory with single, double, and quadruple substitutions [MP4(SDQ)] [16]. These optimizations, and their corresponding frequency calculations, employed the aug-cc-pVTZ basis set. The frequency calculations ensured that all structures were optimized to a true energetic minima on the potential energy surface (PES), and the unscaled thermochemical corrections were used to obtain the zero-point energy (ZPE), enthalpy, and Gibbs free energy. Furthermore, to estimate the energy at the complete basis set limit, we have performed a series of two-point extrapolations on the correlation energy [10, 17–19] In this scheme [Eqs. (1)–(3)], an extrapolated value for the correlation contribution to the total energy is obtained using two consecutive correlation energies, $x - 1$ and x, and is then added to a non-extrapolated Hartree–Fock energy:

$$E_x^{\text{corr}} = E_x^{\text{CCSD(T)}} - E_x^{\text{HF}} \tag{1}$$

$$E_{x-1,x}^{\text{corr}} = \frac{x^3 E_x^{\text{corr}} - (x-1)^3 E_{x-1}^{\text{corr}}}{x^3 - (x-1)^3} \tag{2}$$

$$E_{x-1,x} = E_x^{\text{HF}} + E_{x-1,x}^{\text{corr}}. \tag{3}$$

Additionally, we corrected for basis set superposition error (BSSE) in the HCl and H_2O systems, using the Boys–Bernardi functional counterpoise scheme (fCP) [20].

The W1 model chemistry was developed to be an affordable and accurate method for the determination of thermochemistry [10]. W1 contains one molecule-independent empirical parameter and is based on a series of coupled-cluster calculations, using a variety of robust basis sets. These calculations employ inner shell correlation contributions, scalar relativistic corrections, and two separate two-point extrapolation schemes for both the valence CCSD energy and valence (T) energy contributions.

Obtaining enthalpies and free energies of deprotonation requires the value of H° and G° for the proton. We used the translational energy of $1.5RT$ combined with $PV = RT$ ($H = E + PV$) to obtain a value of $H°$ (H^+) equal to $5/2(RT)$, or 1.48 kcal/mol. We used the Sackur–Tetrode equation (see Ref. [21]) to obtain the entropy, TS(H^+) = 7.76 kcal/mol at 298K and 1 atm pressure, which gives a value (G = H − TS) for G° (H^+) of −6.28 kcal/mol [22, 23]. All values reported in the present work are for a standard state of 1 atm.

Results

Tables I and II compare the model chemistry results for 17 deprotonation reactions with the experimental results reported in the NIST database [6]. Also presented are previously computed thermochemical results using the CBS-QB3 [5], CBS-APNO [4, 24], and G3 [2] model chemistries [7], as well as the B3P86 and PBE1PBE DFT methods [8]. The standard deviation from experiment, σ, is also included in Tables I and II. The experimental values were taken from the NIST database [6]. Selected systems favored small reported error bars, and recently published data [25–40].

Discussion

MODEL CHEMISTRIES

The W1 and CBS-APNO model chemistry results have standard deviations of ~1 kcal/mol relative to

TABLE I

Changes in standard free energies (in kcal/mol) for deprotonation of 17 reactants.*

ΔG° (kcal/mol)		CCSD(T)/aug-cc-pVnZ//MP4(SDQ)/aug-cc-pVTZ										Model chemistry				DFT†	
Reactant	Experimental	T		Q		5		TQ		Q5		CBS-QB3	G3	CBS-APNO	W1	B3P86	PBE1PBE
C_2H_2	370.29 ± 0.41[a]	371.37		370.92		371.08		370.45		371.18		370.33	370.70	370.28	370.61	371.29	370.64
NH_3	396.01 ± 0.20[b]	396.87		396.62		396.61		396.32		396.59		397.40	397.52	396.37	396.40	398.20	398.24
C_2H_4	401.00 ± 0.50[c]	400.70		400.92		401.05		400.93		401.12		400.37	401.06	400.72	400.56	400.90	400.32
CH_2O	386.59 ± 0.31[d]	388.30		387.62		387.70		386.97		387.75		388.84	387.66	387.72	387.54	386.63	386.36
HCl	328.11 ± 0.20[f]	330.40	*326.84*	328.87	*327.32*	330.71	*327.77*	327.33	*327.23*	332.20	*327.80*	326.78	328.75	N/A	328.00	328.19	327.78
CHNO	333.60 ± 0.50[g]	331.49		334.28		333.65		332.92		333.57		334.49	335.51	335.43	335.15	338.18	337.88
CH_4	408.60 ± 0.79[h]	406.48		408.99		408.98		408.91		409.03		410.86	410.99	410.45	411.11	410.28	409.62
CH_3OH	375.0 ± 1.1[i]	374.00		376.75		376.47		376.09		376.65		375.52	376.45	375.81	376.61	374.12	374.05
HNO_2	333.70 ± 0.31[j]	332.09		333.94		333.42		332.82		333.17		331.67	331.93	332.11	333.17	332.90	332.73
H_2O	383.70 ± 0.20[p]	381.97	*382.26*	384.55	*383.26*	384.31	*383.55*	383.86	*383.70*	384.26	*383.76*	385.42	384.75	383.70	384.14	384.77	385.04
CH_3NH_2	395.70 ± 0.69[k]											395.95	396.86	395.88	395.92	395.74	395.64
CH_3NHCH_3	389.20 ± 0.60[k]											387.19	388.45	387.86	388.74	387.14	386.84
$CH_3CH_2NH_2$	391.71 ± 0.69[k]											390.79	391.59	391.19	391.49	391.51	391.17
CH_3CHCH_2	383.80 ± 0.10[l]											383.10	383.86	383.14	383.22	380.81	380.07
C_6H_6	392.90 ± 0.41[m]											391.61	392.46	392.05	392.31	394.39	393.86
HNO_3	317.81 ± 0.20[n]											316.21	317.81	317.94	318.51	318.10	317.83
$C_4H_4O(\alpha)$	382.91 ± 0.20[o]											382.43	382.85	382.79	382.97	384.59	384.32
σ_{top}		1.27	*1.17*	0.76	*0.75*	1.17	*0.76*	0.62	*0.63*	1.59	*0.80*	1.58	1.48	1.20	1.21	2.01	1.89
σ_{bottom}												1.28	0.59	0.74	0.50	1.75	1.94
σ_{total}												1.42	1.17	0.99	0.95	1.84	1.85

* Values in italics have been corrected for BSSE.

† DFT/aug-cc-pVTZ//DFT/aug-cc/pvTZ.

[a][25]; [b][26]; [c][27]; [d][28]; [e][29]; [f][30]; [g][31]; [h][32]; [i][33]; [j][34]; [k][35]; [l][36]; [m][37]; [n][38]; [o][39]; [p][40].

TABLE II

Changes in standard enthalpies (in kcal/mol) for deprotonation of 17 reactants.*

ΔH° (kcal/mol)		CCSD(T)/aug-cc-pVnZ//MP4(SDQ)/aug-cc-pVTZ										Model chemistry				DFT†	
Reactant	Experimental	T		Q		5		TQ		Q5		CBS-QB3	G3	CBS-APNO	W1	B3P86	PBE1PBE
C_2H_2	378.49 ± 0.20[a]	379.38		378.93		379.10		378.47		379.20		378.33	378.70	378.30	378.58	379.17	378.54
NH_3	403.39 ± 0.10[b]	404.38		404.14		404.13		403.84		404.11		404.95	405.07	403.91	403.93	405.11	405.14
C_2H_4	409.39 ± 0.60[c]	409.09		409.30		409.43		409.31		409.50		408.76	409.45	409.12	408.95	409.29	408.71
CH_2O	394.50 ± 0.10[d]	396.23		395.55		395.64		394.90		395.68		397.18	395.63	395.73	395.54	394.63	394.35
HCl	333.39 ± 0.10[f]	335.78	*332.22*	334.25	*332.70*	336.09	*333.15*	332.71	*332.61*	337.58	*333.18*	332.15	334.14	N/A	333.37	333.56	333.15
CHNO	340.01 ± 0.41[g]	342.61		341.98		341.98		341.26		341.90		340.88	341.88	341.79	341.55	344.59	344.28
CH_4	416.71 ± 0.69[h]	417.28		417.27		417.29		417.20		417.32		419.13	419.26	418.73	417.91	419.21	418.55
CH_3OH	381.7 ± 1.0[i]	383.24		382.96		383.08		382.58		383.13		382.69	383.61	382.99	383.09	381.32	381.27
HNO_2	340.20 ± 0.20[f]	340.95		340.44		340.37		339.84		340.19		339.11	339.39	339.58	340.18	339.97	339.81
H_2O	390.3 ± 0.20[p]	391.15	*388.86*	390.90	*389.86*	390.92	*390.15*	390.46	*390.29*	390.85	*390.35*	392.03	391.37	390.31	390.74	391.37	391.64
CH_3NH_2	403.20 ± 0.79[k]											403.38	404.30	403.34	403.37	403.18	403.08
CH_3NHCH_3	396.39 ± 0.91[k]											394.50	395.79	395.19	395.62	394.07	393.78
$CH_3CH_2NH_2$	399.4 ± 1.1[k]											398.24	399.03	398.65	398.97	398.97	398.64
CH_3CHCH_2	391.11 ± 0.31[l]											391.38	392.02	391.02	390.81	388.56	387.98
C_6H_6	401.70 ± 0.50[m]											400.51	401.27	400.87	401.18	402.63	402.09
HNO_3	324.50 ± 0.20[n]											323.57	324.17	324.31	324.81	325.09	324.84
$C_4H_4O(\alpha)$	391.11 ± 0.41[o]											390.66	391.05	390.98	391.18	392.81	392.52
σ_{top}		1.50	*1.40*	1.07	*1.04*	1.35	*1.03*	0.91	*0.92*	1.69	*1.03*	1.75	1.57	1.32	1.03	2.01	1.89
σ_{bottom}												1.12	0.69	0.68	0.46	1.64	1.80
σ_{total}												1.50	1.27	1.07	0.83	1.81	1.80

* Values in italics have been corrected for BSSE.

† DFT/aug-cc-pVTZ//DFT/aug-cc/pvTZ.

[a][25]; [b][26]; [c][27]; [d][28]; [e][29]; [f][30]; [g][31]; [h][32]; [i][33]; [j][34]; [k][35]; [l][36]; [m][37]; [n][38]; [o][39]; [p][40].

the experimental values, while the computationally less expensive G3 method yields standard deviations of 1.27 and 1.17 kcal/mol for enthalpy and free energy changes. The CBS-QB3 method, which uses DFT geometries, has standard deviations of 1.5 and 1.4 kcal/mol for enthalpy and free energy changes. W1 is the only method that has a higher standard deviation for free energy changes than enthalpy changes, suggesting that the scaling of the B3LYP/cc-pVTZ frequencies within that method is not optimal for free energies.

A comparison of the model chemistry results to the CCSD(T) results is only possible for the smallest 10 molecules at the top of Table I because of the expense of the coupled clusters calculation scheme. The standard deviations for the CBS-QB3, G3, CBS-APNO, and W1 results of the 10 atom set, σ_{top}, are 1.8, 1.6, 1.3, and 1.0 for enthalpies and 1.6, 1.5, 1.2, and 1.2 for free energies. This is interesting for two reasons, the first being that the coupled cluster method using a quadruple-zeta basis set is significantly more accurate than the model chemistries. The second, which may be more important for users of these methods, is that the model chemistries are more accurate for the seven largest molecules in this test set. As shown by σ_{bottom}, the G3, CBS-APNO, and W1 model chemistries are better for the larger molecules which are extremely costly to calculate with the CCSD(T) method.

The G3 and CBS-APNO methods model the geometries and energies of hydrogen-bonded water and ion-water clusters quite accurately [41–48], and this deprotonation study adds to the body of evidence that shows the reliability of these two model chemistries. Furthermore, the W1 method, which is computationally more expensive, is also highly accurate. These three model chemistries are more reliable than DFT methods that use the aug-cc-pVTZ basis set for energy calculations [8], and are comparable with the best CCSD(T) results displayed in Tables I and II.

COUPLED CLUSTER CALCULATIONS

Tables I and II show that the CCSD(T)/aug-cc-pV*n*Z method has the best overall agreement with experiment when n = Q. While collectively the n = 5 results are worse than n = Q, when the HCl outlier is removed the accuracy of the two methods are roughly equivalent. Because the errors associated with the n = T and Q calculations systematically decrease from T to Q, the TQ two-point extrapolation results are in excellent agreement with experiment. Conversely, there are no systematic changes in the errors between the n = Q and 5 calculations, which is reflected in the extrapolation results. Ignoring the HCl outlier, the calculated values for enthalpy and free energy of reaction are virtually unchanged as the basis is expanded from n = Q to 5. Excluding HCl, recalculation of the standard deviation from experiment for the Q5 extrapolation results in σ falling from 1.59 to 0.86 and from 1.69 to 1.10 for free energy and enthalpy respectively, a nearly equivalent result to the TQ extrapolation. We feel that ignoring the HCl outlier is justifiable, particularly within the context of our extrapolation scheme because the error in the HCl calculated correlation energy increases as the basis set expands, and our two-point extrapolation scheme amplifies this error. The failure of the quintuple-zeta HCl calculation inspired us to perform a very limited set of fCP corrections. We investigated BSSE errors for the aforementioned HCl system, as well as the H_2O system, which enjoyed favorable agreement with experiment. As shown by the values in italics in Tables I and II, correction for BSSE brings the values for the n = T, Q, and 5 basis sets into better agreement, and the extrapolated TQ and Q5 values are now both in excellent agreement with experiment. Because of the extreme expense of these calculations, we were unable to perform fCP corrections for all ten of the systems studied with coupled cluster theory, and thus we are unable to determine whether this is a fortuitous result or a general trend. However, these results do suggest that even high correlation methods executed with robust basis sets still benefit from fortuitous error cancellation. Feller et al. [49] recently published a careful study of the sources of error for electronic structure calculations on small systems. They used correlation consistent basis sets ranging up to the n = 10 size to show that coupled cluster calculations slowly converge on the total energy as n increases. The raw CCSD(T)/cc-pv10Z energy of the Ne (^{1}S) atom remains 0.6 kcal/mol above the estimated complete basis set limit [49]. Because energy differences converge more rapidly than raw energies, the bond dissociation energy of N_2 is converged to 0.5 kcal/mol at the CCSD(T)/cc-pV7Z level [49]. Using simple extrapolation schemes effectively improves the computed properties by an equivalent of increasing the basis set size by three or more cardinal numbers [49]. This suggests that a single-point calculation at the n = 7 or n = 8 level yields results similar to an extrapolated TQ or Q5 scheme. Additionally, the CCSD(T) method often

produces results that are in better agreement with full configuration interaction than the more expensive CCSDT method [49]. Thus, we advocate the use of extrapolated TQ values, because of their relatively inexpensive, yet high accuracy predictions. The G3, CBS-APNO, and W1 model chemistry methods appear to be the more efficient route when calculating the thermodynamics of deprotonation for small organic molecules.

Interestingly, the coupled clusters and model chemistry results both consistently disagree with certain experimental results. The NIST values for $\Delta G°$ and $\Delta H°$ have been revised along with the error bars for the experimental deprotonation of methanol since our first use of this data set [7]. The original error bars of 0.6 and 0.4 kcal/mol in the NIST database in 2000 have been revised to 1.1 and 1.0 kcal/mol, and the results here suggest that the error bars are even larger, as the value for $\Delta G°$ should be closer to 376 rather than 375 kcal/mol.

Conclusion

The model chemistry methods G3, CBS-APNO, and W1 are relatively fast and efficient methods for obtaining deprotonation enthalpies and free energies. Calculations at the CCSD(T)/aug-cc-pVQZ level are very accurate for the 17-molecule test set. Results are worse for a coupled cluster energy calculation using the aug-cc-pV5Z data set, stemming from either better cancellation of errors for the quadruple-zeta basis set or from the effects of BSSE. Any of the model chemistry methods or the CCSD(T)/aug-cc-pVQZ method can serve as a valuable check on the accuracy of experimental data reported in the NIST database.

ACKNOWLEDGMENTS

The authors thank Greg Nizialek and Christy House for preliminary work, and to NSF for support of the MERCURY high-performance computer consortium (http://mercury.chem.hamilton.edu).

References

1. Curtiss, L. A.; Raghavachari, K.; Trucks, G. W.; Pople, J. A. J Chem Phys 1991, 94, 7221.
2. Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Rassolov, V.; Pople, J. A. J Chem Phys 1998, 109, 7764.
3. Petersson, G. A.; Tensfeldt, T. G.; Montgomery, J. A., Jr. J Chem Phys 1991, 94, 6091.
4. Ochterski, J. W.; Petersson, G. A.; Montgomery, J. A., Jr. J Chem Phys 1996, 104, 2598.
5. Montgomery, J. A., Jr.; Frisch, M. J.; Ochterski, J. W.; Petersson, G. A. J Chem Phys 1999, 110, 2822.
6. Bartmess, J. E. In NIST Chemistry WebBook; NIST Standard Reference Database No. 69 June 2005 Release; Mallard, W. G.; Linstrom, P. J., Eds.; National Institute of Standards and Technology: Gaithersburg, MD, 2000.
7. Pokon, E. K.; Liptak, M. D.; Feldgus, S.; Shields, G. C. J Phys Chem A 2001, 105, 10483.
8. Liptak, M. D.; Shields, G. C. Int J Quantum Chem 2005, 105, 580.
9. Pople, J. A.; Head-Gordon, M.; Raghavachari, K. J Chem Phys 1987, 87, 5968.
10. Martin, J. M. L.; de Oliveira, G. J Chem Phys 1999, 111, 1843.
11. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. Gaussian 03; Gaussian: Pittsburgh, PA, 2003.
12. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennuci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B. G.; Chem, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. Gaussian 98; Gaussian: Pittsburgh, PA, 1998.
13. Dunning, T. H., Jr. J Chem Phys 1989, 90, 1007.
14. Woon, D. E.; Dunning, T., Jr. H. J Chem Phys 1993, 98, 1358.
15. Woon, D. E.; Dunning, T. H., Jr. J Chem Phys 1995, 103, 4572.
16. Krishnan, R.; Pople, J. A. Int J Quantum Chem 1978, 14, 91.
17. Bak, K. L.; Jorgensen, P.; Olsen, J.; Helgaker, T.; Klopper, W. J Chem Phys 2000, 112, 9229.
18. Halkier, A.; Helgaker, T.; Jorgensen, P.; Klopper, W.; Koch, H.; Olsen, J.; Wilson, A. K. Chem Phys Lett 1998, 286, 243.

19. Helgaker, T.; Klopper, W.; Koch, H.; Noga, J. J Chem Phys 1997, 106, 9639.
20. Boys, S. F.; Bernardi, R. Mol Phys 1970, 19, 553.
21. McQuarrie, D. M. Statistical Mechanics; Harper & Row: New York, 1970.
22. Liptak, M. D.; Shields, G. C. J Am Chem Soc 2001, 123, 7314.
23. Tawa, G. J.; Topol, I. A.; Burt, S. K.; Caldwell, R. A.; Rashin, A. A. J Chem Phys 1998, 109, 4852.
24. Petersson, G. A.; Allaham, M. A. J Chem Phys 1991, 94, 6081.
25. Ervin, K. M.; Lineberger, W. C. J Phys Chem 1991, 95, 1167.
26. Wickham-Jones, C. T.; Ervin, K. M.; Ellison, G. B.; Lineberger, W. C. J Chem Phys 1989, 91, 2762.
27. Ervin, K. M.; Gronert, S.; Barlow, S. E.; Gilles, M. K.; Harrison, A. G.; Bierbaum, V. M.; DePuy, C. H.; Lineberger, W. C.; Ellison, G. B. J Am Chem Soc 1990, 112, 5750.
28. Murray, K. K.; Miller, T. M.; Leopold, D. G.; Lineberger, W. C. J Chem Phys 1986, 84, 2520.
29. Kim, E. H.; Bradforth, S. E.; Arnold, D. W.; Metz, R. B.; Neumark, D. M. J Chem Phys 1995, 103, 7801.
30. Trainham, R.; Fletcher, G. D.; Larson, D. J. J Phys B: At Mol Phys 1987, 20, L777–L784.
31. Bradforth, S. E.; Kim, E. H.; Arnold, D. W.; Neumark, D. M. J Chem Phys 1993, 98, 800.
32. Ellison, G. B.; Engelking, P. C.; Lineberger, W. C. J Am Chem Soc 1978, 100, 2556.
33. Ramond, T. M.; Davico, G. E.; Schwartz, R. L.; Lineberger, W. C. J Chem Phys 2000, 112, 1158.
34. Ervin, K. M.; Ho, J.; Lineberger, W. C. J Phys Chem 1988, 92, 5405.
35. Mackay, G. I.; Hemsworth, R. S.; Bohme, D. K. Can J Chem 1976, 54, 1624.
36. Ellison, G. B.; Davico, G. E.; Bierbaum, V. M.; DePuy, C. H. Int J Mass Spectrom Ion Proc 1996, 156, 109.
37. Davico, G. E.; Bierbaum, V. M.; DePuy, C. H.; Ellison, G. B.; Squires, R. R. J Am Chem Soc 1995, 117, 2590.
38. Davidson, J. A.; Fehsenfeld, F. C.; Howard, C. J. Int J Chem Kinet 1977, 9, 17.
39. Clifford, E. P.; Wenthold, P. G.; Lineberger, W. C.; Ellison, G. B.; Wang, C. X.; Grabowski, J. J.; Vila, F.; Jordan, K. D. J Chem Soc Perkin Trans II 1998, 1015.
40. Schulz, P. A.; Mead, R. D.; Jones, P. L.; Lineberger, W. C. J Chem Phys 1982, 77, 1153.
41. Dunn, M. E.; Pokon, E. K.; Shields, G. C. Int J Quantum Chem 2004, 100, 1065.
42. Dunn, M. E.; Pokon, E. K.; Shields, G. C. J Am Chem Soc 2004, 126, 2647.
43. Pickard, F. C.; Pokon, E. K.; Liptak, M. D.; Shields, G. C. J Chem Phys 2005, 122.
44. Pickard, F. C.; Dunn, M. E.; Shields, G. C. J Phys Chem A 2005, 109, 4905.
45. Day, M. B.; Kirschner, K. N.; Shields, G. C. Int J Quantum Chem 2005, 102, 565.
46. Day, M. B.; Kirschner, K. N.; Shields, G. C. J Phys Chem A 2005, 109, 6773.
47. Dunn, M. E.; Evans, T. M.; Kirschner, K. N.; Shields, G. C. J Phys Chem A 2006, 110, 303.
48. Alongi, K. S.; Dibble, T. S.; Shields, G. C.; Kirschner, K. N. J Phys Chem A 2006, 110, 3686.
49. Feller, D.; Peterson, K. A.; Crawford, T. D. J Chem Phys 2006, 124, 054107.

Exploration of the Potential Energy Surfaces, Prediction of Atmospheric Concentrations, and Prediction of Vibrational Spectra for the $HO_2\cdots(H_2O)_n$ ($n = 1-2$) Hydrogen Bonded Complexes

Kristin S. Alongi,[†] Theodore S. Dibble,*[,‡] George C. Shields,*[,†] and Karl N. Kirschner*[,†]

Department of Chemistry, Hamilton College, Clinton, New York 13323, and Chemistry Department, SUNY-ESF, Syracuse, New York 13210

Received: December 8, 2005; In Final Form: January 25, 2006

The hydroperoxy radical (HO_2) plays a critical role in Earth's atmospheric chemistry as a component of many important reactions. The self-reaction of hydroperoxy radicals in the gas phase is strongly affected by the presence of water vapor. In this work, we explore the potential energy surfaces of hydroperoxy radicals hydrogen bonded to one or two water molecules, and predict atmospheric concentrations and vibrational spectra of these complexes. We predict that when the HO_2 concentration is on the order of 10^8 molecules·cm^{-3} at 298 K, that the number of $HO_2\cdots H_2O$ complexes is on the order of 10^7 molecules·cm^{-3} and the number of $HO_2\cdots(H_2O)_2$ complexes is on the order of 10^6 molecules·cm^{-3}. Using the computed abundance of $HO_2\cdots H_2O$, we predict that, at 298 K, the bimolecular rate constant for $HO_2\cdots H_2O + HO_2$ is about 10 times that for $HO_2 + HO_2$.

Introduction

The hydroperoxy radical (HO_2) is a molecule that is of interest in many fields, including environmental chemistry, astrochemistry, and biochemistry. It plays a significant role in Earth's atmospheric chemistry as a component of several important gas-phase reactions. The formation of the hydroperoxy radical stems primarily from the OH radical-initiated degradation of organic compounds, and the subsequent photolysis of aldehydes:[1]

$$H_2CO(g) + h\nu\ (\lambda \leq 370\ \text{nm}) \rightarrow HCO(g) + H(g) \quad (1)$$

$$H(g) + O_2 \rightarrow HO_2(g) \quad (2)$$

$$HCO(g) + O_2 \rightarrow HO_2(g) + CO(g) \quad (3)$$

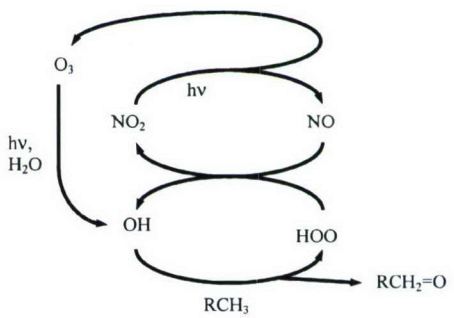
The self-reaction of two hydroperoxy radicals leads to the production of molecular oxygen and hydrogen peroxide:[2]

$$HO_2(g) + HO_2(g) \rightarrow HOOH(g) + O_2(g) \quad (4)$$

Once formed, hydrogen peroxide rapidly enters aqueous aerosols where it oxidizes sulfur dioxide to sulfuric acid. The reaction of hydroperoxy radical with nitric oxide forms NO_2, whose photolysis leads to formation of ozone as shown in Scheme 1.

The hydroperoxy radical also has a very strong affinity for water. The atmospheric importance of this fact came to light 3 decades ago, when it was determined that the rate of hydroperoxy radical's self-reaction in the gas phase was strongly affected by the presence of water vapor.[2−7] This was widely suspected to result from the formation of an $HO_2\cdots H_2O$ complex, and led to one of the first ab initio studies of this hydrogen bonded complex.[8] Astronomers are also interested in understanding the interaction between hydroperoxy and water molecules as related to their interpretation of data obtained from icy surfaces in space. Radiolysis of icy surfaces on comets, interstellar dust, and satellites (e.g., Jupiter's moon Europa) could result in the chemical processes seen in hydroperoxy radical's self-reaction or its reaction with water.[9,10] The hydroperoxy radical is also believed to be important in several biological processes, where it has ample opportunity for forming weak interactions with water molecules.[11] Thus, understanding the structure and energetics of the $HO_2\cdots H_2O$ complex is fundamentally important for several fields of study.

SCHEME 1: Catalytic Cycles in HO_x and NO_x Producing Ozone While Oxidizing Volatile Organic Compounds



There has been an indirect determination of the equilibrium constant for formation of the $HO_2\cdots H_2O$ complex[12] and an attempt to obtain its spectrum.[13] However, no interpretable spectrum has been obtained for this complex. Recently, there have been several quantum-mechanical studies of a single conformer of the $HO_2\cdots H_2O$ complex.[14−16] Because of the role water has in hydroperoxy radical reactions, gaining a detailed energetic and structural understanding of possible $HO_2\cdots H_2O$ and $HO_2\cdots(H_2O)_2$ complexes will give additional insight into the role this hydrogen bonded system plays in atmospheric chemistry and other phenomena.

Using CCSD//MP2 calculations with a variety of basis sets, the equilibrium constant (standard state of molecule·cm^3) for reaction 5

[†] Hamilton College.
[‡] SUNY-ESF.

10.1021/jp057165k CCC: $33.50
Published on Web 02/22/2006

$$HO_2(g) + H_2O(g) \rightarrow HO_2\cdots H_2O(g) \quad (5)$$

at room temperature has been estimated to be between 4×10^{-18} and 1.3×10^{-21}.[14,17] Experimentally, Kanno and colleagues determined the reaction's equilibrium constant K_c to be $(5.2 \pm 3.2) \times 10^{-19}$ at 297 K, which leads to a concentration ratio of $[HO_2\cdots H_2O]/[HO_2]$ of 0.19 ± 0.11 at 297 K and 50% relative humidity.[12]

Hydroperoxy radical's self-reaction has been the subject of three computational studies to date,[15,18,19] two of which have examined the catalytic role of water. In addition to the $HO_2\cdots H_2O$ dimer complex, the $HO_2\cdots(H_2O)_2$ trimer is of interest. The trimer may be atmospherically relevant, and serves as a stepping stone to modeling bulk or surface hydration of HO_2.[16,20] As a model for HO_2 interacting with cloud droplets, Shi and co-workers performed quantum-mechanical calculations on the $HO_2\cdots(H_2O)_{20}$ complex.[20] They proposed that hydroperoxy radical reactions may occur on the surface and in the interior of a cloud droplet. Complexes and clusters of water with oxidants, including the hydroperoxy radical, is the subject of two recent reviews.[21,22] Our goal is to obtain all of the potential configurations of $HO_2\cdots H_2O$ and $HO_2\cdots(H_2O)_2$ complexes, and determine their energies and relative abundances in the lowermost troposphere. This is a fundamental step for obtaining a better understanding of water's role in the self-reaction of the hydroperoxy radical and its gas- and aqueous-phase chemistry.

Method

The initial $HO_2\cdots H_2O$ configurations were built using the SPARTAN[23] software by placing a water molecule about the hydroperoxy radical at three different locations. In two of these configurations, each radical oxygen atom accepted a hydrogen bond from the water, while the third configuration had the water accepting a hydrogen bond from the hydroperoxy radical. Additional configurations were generated using a 12-fold rotation around the hydrogen bond of each initial structure, and each of these configurations was optimized with PM3. The resulting conformers were grouped based on dihedrals of matching signs (positive or negative) and similar values (i.e., 0−8, 8−25, 25−90, 90−120, 120+). Previous work on small water clusters has shown that PM3 geometry optimizations of conformers with similar dihedral angles converge on a common minimum.[24] The groups were not based on energies because most conformations from each search possessed PM3 heat of formation energies within 3 kcal·mol^{-1} of each other. We performed a Hartree−Fock (HF) 6-31G* optimization using the Gaussian03 program[25] on the lowest energy member from each of these groups. The resulting HF energy values and structures were then compared to identify unique complexes. Some of the structures picked from different groups optimized to similar structures, and in total, only two unique structures (dimers A and C) were located. A third dimer configuration (dimer B) was found by enforcing C_s symmetry on the molecule. Without C_s symmetry, this configuration quickly minimizes to dimer C in Figure 1. Finally, Gibbs free energies were obtained using the Gaussian (G3) model chemistry.[26] No further corrections have been made for basis set superposition error, as the method has an inherent correction for basis set artifacts.[27−29]

We started our search for the $HO_2\cdots(H_2O)_2$ complexes by bonding another water molecule to the original $HO_2\cdots H_2O$ complexes in various locations. This resulted in 13 new starting structures. For the exploration of the conformations of these structures, over one hundred additional structures were generated using either a 12-fold or an 8-fold rotation around the hydrogen bond, prior to PM3 optimization of each conformer. The conformers were grouped based on dihedral angles, optimized with HF/6-31G* calculations, grouped based on energies, and the remaining conformers were input to the G3 model chemistry. This approach generated four $HO_2\cdots(H_2O)_2$ complexes. We performed frequency calculations on all complexes using the HF/6-31G* level of theory, and computed a Boltzmann distribution using our G3 model chemistry free energies. All complexes reported have been verified to be stable minima on the HF/6-31G* potential energy surface. All calculations involving hydroperoxy radical were done using spin-unrestricted wave functions.

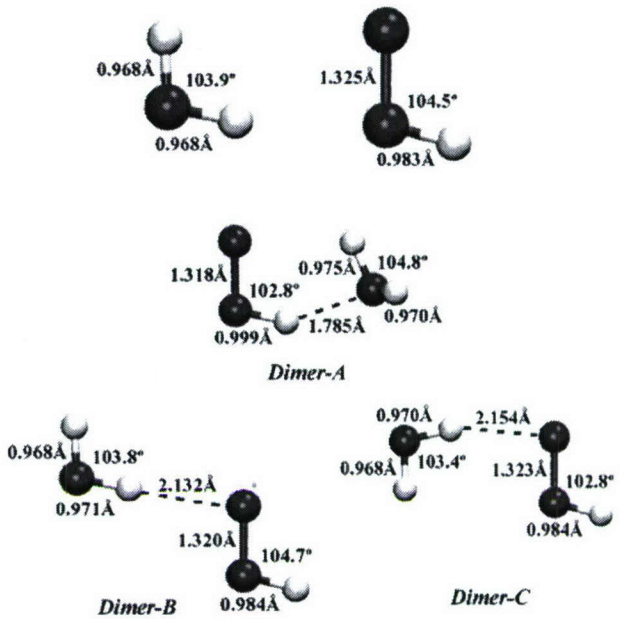


Figure 1. Molecular structures of HO_2, H_2O, and the three $HO_2\cdots H_2O$ dimers determined at the MP2/6-31G* level of theory. Key interatomic distances (angstroms) and angles (degrees) are given.

Results and Discussion

Structures. We found three unique hydrogen bonded complexes of the $HO_2\cdots H_2O$ heterodimer, including the one previously described in the literature where the hydroperoxy radical molecule is the hydrogen bond donor (dimer A in Figure 1). The two new dimer configurations (dimers B and C) both involve the hydroperoxy radical acting as a hydrogen bond acceptor.

Dimer A belongs to the C_1 point group and has an enantiomer. Our MP2(full)/6-31G* optimized value for the hydrogen bond length is 1.785 Å, only 0.004 Å shorter than the MP2/6-311++G(2df,2pd) value of Aloisio and Francisco.[14] The excellent agreement between MP2(full)/6-31G* and MP2/6-311++G(2df,2pd) results allows us to have confidence in the geometries of dimers B and C. Both dimers B and C possess C_s symmetry. Dimer B has a hydrogen bond length of 2.132 Å, indicating weaker attractive forces involved in this dimer formation in comparison to dimer A. Breaking the C_s symmetry in dimer B followed by an optimization forms dimer C, the third and least stable dimer (rotating dimer B's water by 180° results in the formation of dimer C, recapturing C_s symmetry). dimer C has a 2.154 Å hydrogen bond, which is the longest hydrogen bond distance seen in any of the dimer or trimer complexes.

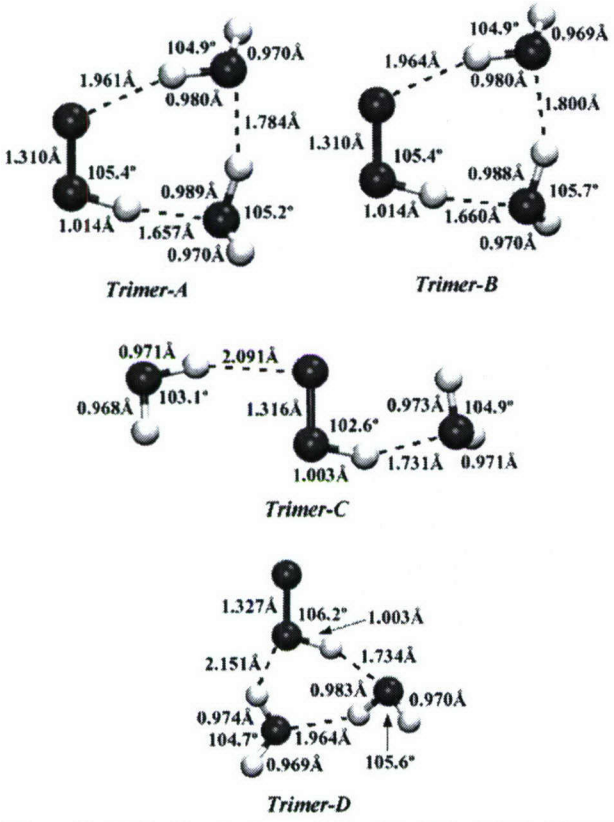


Figure 2. Molecular structures of the four $HO_2\cdots(H_2O)_2$ trimers determined at the MP2/6-31G* level of theory. Key interatomic distances (angstroms) and angles (degrees) are given.

We found four unique configurations for the $HO_2\cdots(H_2O)_2$ trimer complex, all of which have the hydroperoxy radical acting as a hydrogen bond donor and acceptor, as seen in Figure 2. Each trimer possesses C_1 symmetry and has an enantiomer, with trimers A, B, and D existing as cyclic clusters. The only difference between trimers A and B is the orientation of the hydrogen on the water molecules. In these trimers, the cyclic structure allows both oxygens of the hydroperoxy radical to participate in hydrogen bonding with a water, one as a hydrogen bond donor and one as an acceptor. By contrast, in trimer C this is achieved with an extended structure; thus trimer C is the only cluster where the two water molecules do not form a hydrogen bond between each other. Trimer D is unique as the hydroperoxy radical's hydroxyl oxygen acts both as a hydrogen-bond donor and acceptor, while its other oxygen (the radical center) does not participate in any noncovalent interactions. An interesting observation about the trimer complexes is that the hydrogen bonds formed by hydroperoxy radical donating hydrogen are shorter than that observed in dimer A, which has a similar bonding motif. This is an indication that the inclusion of the second water alters the electronic configuration about dimer A, allowing for a more energetically favorable interaction to occur. The cyclic trimers have the same cooperativity interactions as cyclic water clusters that have hydrogen bonds that all donate in the same direction.[24,27,30–40] The most stable $HO_2\cdots(H_2O)_2$ clusters can be constructed simply by replacing one of the water molecules in the most stable water trimers[24] with the radical.

Thermochemistry. As can be seen in Table 1, dimer A is significantly more stable than the two other conformers. For dimer A Aloisio and Francisco reported a CCSD(T)/6-311++G-(2df,2p)//B3LYP/6-311++G(2df,2p) electronic energy for reaction 5, $\Delta E_{elec}(0\ K)$, of -9.4 kcal·mol^{-1} and after correcting for zero-point vibrational energy, ΔE_{ZPVE}, the energy becomes -6.9 kcal·mol^{-1}.[14] These values are in excellent agreement with our G3 $\Delta E_{elec}(0\ K)$ and ΔE_{ZPVE} values of -9.1 and -7.1 kcal·mol^{-1}. Including thermal and entropic effects results in dimer A having a enthalpy, $\Delta H°(298\ K)$, and free energy for reaction 5, $\Delta G°(298\ K)$, of -7.4 and -0.5 kcal·mol^{-1}. Dimers B and C have a significantly more positive free energy for reaction 5, with values of 2.4 and 3.1 kcal·mol^{-1}. The fact that dimer A consists of two enantiomers further lowers the effective free energy by RT ln(2), for an effective $\Delta G°(298\ K)$ of -0.94 kcal·mol^{-1}. A Boltzmann calculation, taking into account the enantiomer of dimer A, predicts the relative abundances at 298.15 K of dimers A, B, and C to be 99.5%, 0.4%, and 0.1%. Dimers of hydroperoxy radical and a water molecule will be dominated by dimer A, where the hydroperoxy radical acts as the hydrogen bond donor.

The calculated Gibbs free energy for the reaction forming the two enantiomers of dimer A is -0.9 kcal·mol^{-1}, which is in good agreement with the experimental value of -1.5 ($+0.6/-0.3$) kcal·mol^{-1} from Kanno and co-workers as determined from their reported equilibrium constant.[12] The relative instability of dimers B and C means that they contribute negligibly to the overall dimer population.

Let us now consider the thermodynamics of the formation of the trimers:

$$HO_2(g) + 2H_2O(g) \rightarrow HO_2\cdots(H_2O)_2(g) \qquad (6)$$

Table 1 contains the energies and the Boltzmann distribution calculation at 298 K for the trimers. As expected based on the strong similarities of their structures, trimers A and B possess nearly the same energy, with $\Delta G°$ (298 K) for reaction 6 of -0.47 and -0.39 kcal·mol^{-1}, respectively. Trimers A and B are also the most stable trimers, and are responsible for 99% of the trimer population. Hydrogen bonding in these two trimers occurs in a three-membered ring, with each molecule/radical acting as both a hydrogen bond donor and acceptor. This cyclic motif has been observed in water clusters,[27,35] and has been attributed to the enhanced cooperativity of these cyclic structures.[37]

We now consider the abundance of these dimers and trimers in the atmosphere at 298 K. This depends on the K_p values for reactions 4 and 5, the abundance of water vapor, and the abundance of hydroperoxy radical:

$$K_{p,dimer} = P_{complex}/P_{HO_2}P_{H_2O} \qquad (7a)$$

$$K_{p,trimer} = P_{complex}/P_{HO_2}(P_{H_2O})^2 \qquad (7b)$$

where P_i are expressed in atmospheres. Since there are multiple dimers and trimers, we compute the ratio $P_{complex,i}/P_{HO_2}$ for each complex i, and determine the fraction of the total HO_2 present as each complex:

$$[P_{complex,i}/P_{HO_2}]/\{1 + \textstyle\sum_i [P_{complex,i}/P_{HO_2}]\} \qquad (8)$$

Values of K_p calculated from $\Delta G°(298\ K)$ and accounting for the presence of enantiomers are listed in Table 2. HO_2 concentrations vary enormously in time and space, but often reach 5×10^8 molecules·cm^{-3} in photochemically active regions of the lower troposphere. Present detection methods[41] probably detect $HO_2\cdots H_2O$ and $HO_2\cdots(H_2O)_2$ with roughly the same efficiency as monomer, so we will take 5×10^8 molecules·cm^{-3}

TABLE 1: Changes in G3 Electronic Energies, Enthalpies, Entropies, Gibbs Free Energies, and Boltzmann Distribution for Formation of $HO_2\cdots(H_2O)_n$ for $n = 1$ and 2[a]

$HO_2\cdots(H_2O)_n$	ΔE_{elec}(0 K)	ΔE_{ZPVE}(0 K)	$\Delta H°$(298 K)	$\Delta G°$(298 K)	$\Delta S°$ (cal·mol^{-1}· K^{-1})	distribution (298 K)
$n = 1$						
dimer A	−9.14	−7.06	−7.42	−0.53	76.45	99.5%
dimer B	−3.54	−2.38	−2.15	2.37	84.42	0.4%
dimer C	−3.43	−2.22	−2.04	3.12	82.25	0.1%
$n = 2$						
trimer A	−20.64	−15.80	−16.98	−0.47	89.20	52.5%
trimer B	−20.08	−15.46	−16.52	−0.39	90.45	45.9%
trimer C	−13.53	−10.36	−10.41	1.68	104.04	1.4%
trimer D	−15.71	−11.44	−12.27	2.91	93.66	0.2%

[a] All energies reported are based upon a standard state of 1 atm and have units of kcal·mol^{-1}.

TABLE 2: Equilibrium Constant (Standard State of 1 atm) and Population of Free HO_2 and $HO_2\cdots(H_2O)_n$ for $n = 1$ and 2 at 100% Relative Humidity for Assumed Total [HO_2] = 5×10^8 molecules·cm^{-3} at a Temperature of 298.15 K[a]

species	K_p (1 atm)	N (molecules·cm^{-3})
HO_2 monomer		4.3×10^8
$n = 1$		
dimer A (e)	2.4 (13 ± 8)[b]	6.6×10^7
dimer B	0.018	2.5×10^5
dimer C	0.052	6.9×10^4
n = 2		
trimer A (e)	4.4	1.9×10^6
trimer B (e)	3.9	1.6×10^6
trimer C (e)	0.12	4.9×10^4
trimer D (e)	0.015	6.2×10^3

[a] Equilibrium constants are increased by ln 2 for species with two enantiomers, where parenthesis (e) indicates an enantiomeric pair. [b] Reference 12.

to be the concentration of HO_2 (as monomer and in all clusters), and assume that only dimers and trimers contribute significantly to the HO_2 cluster population. For this calculation we will assume a relative humidity (RH) of 100% ($P_{H_2O} = 0.03125$ atm).

Results of these calculations are listed in Table 2. We predict that dimer A will have a concentration of 6.6×10^7 molecules·cm^{-3}, and dimers B and C will have equilibrium concentrations 2−3 orders of magnitude lower. Trimers A and B will have concentrations of 1.9×10^6 and 1.6×10^6 molecules·cm^{-3}, while trimers C and D will have concentrations 1.5−2.5 orders of magnitude lower. Since the hydroperoxy radical concentration is much less than that of water vapor, the computed dimer and trimer concentrations scale linearly with the assumed total HO_2 concentration. Conversely, dimer and trimer concentrations do not scale linearly with RH; decrease of RH to 50% reduces dimer concentrations by 44% and trimer concentrations by 73%.

As noted in the Introduction, one motivation for this study was an interest in how water vapor enhances the self-reaction of HO_2 (reaction 4). The observed effect is roughly linear with water concentration. One commonly used recommendation[7] is that the bimolecular rate constant of reaction 4 is increased by a multiplicative factor:

$$k(T, [H_2O]) = k(T, [H_2O] = 0)\{1 + 1.4 \times 10^{-21}[H_2O]e^{+2200/T}\} \quad (9)$$

where $[H_2O]$ is in molecules·cm^{-3}. It is widely assumed that the increase in rate constant is due to the reaction:

$$HO_2\cdots H_2O(g) + HO_2(g) \rightarrow HOOH(g) + H_2O(g) + O_2(g) \quad (10)$$

If we make the simple assumption that substituting dimer A for HO_2 monomer increases the reaction rate by an enhancement factor, E, due to the fraction, f, of dimer A present, then at 298.15 K eq 9 is related to E and f by

$$\{1 + 1.4 \times 10^{-21}[H_2O]e^{+2200/298.15}\} = (1 - f) + fE \quad (11)$$

The formation of dimer A reduces the concentration of HO_2 monomer by a fraction f, thereby causing a corresponding decrease in the rate of the self-reaction of HO_2 monomer. The $(1 - f)$ term on the right-hand side corresponds to this fractional decrease. We then solve for E using the range of RH used in experiment (0−55%). Using the values of K_p computed here to determine f, we obtain $E = 10.5$. Using Kanno's value of $K_p = 13 \pm 8$, we compute an enhancement factor of 2−10 (4 [+6, −2]).

Vibrational Spectra. The calculated and experimental frequencies for the water and hydroperoxy monomers, and the three dimer complexes are given in Table 3. The scaled HF/6-31G* frequencies for the water monomer agree very well with the experimental frequencies, with a maximum error of 36 cm^{-1}. The frequency of the υ_1 mode of the hydroperoxy monomer should be diagnostic of hydrogen bond donation by HO_2, so it would be helpful to be able to rely on the computed frequencies of this mode in the complexes. Unfortunately, the scaled frequency we obtain at the HF/6-31G* level overestimates the experimental frequency by 172 cm^{-1}.

The experimental OH stretching frequency of the HO_2 molecule in the dimer complex occurs as a strong infrared peak at 3236 cm^{-1}, redshifted by 177 cm^{-1} with respect to hydroperoxy monomer (both in an Ar matrix).[42] Computations indicate a redshift of 101 cm^{-1} for dimer A, but only 3 and 1 cm^{-1} for dimers B and C. Clearly, the redshift seen in the experimental spectrum of the dimer is only consistent with dimer A. Note that the Ar matrix only redshifts υ_1 of hydroperoxy monomer by ~20 cm^{-1} with respect to the gas-phase value.[43,44]

The scaled vibrational frequencies for the four trimers are reported in Table 4 along with their relative infrared and Raman intensities. The "a" and "d" by the mode labels in Table 4 indicate whether the species involved is acting as a hydrogen bond acceptor or donor in that mode. The intermolecular frequencies all exist below 1000 cm^{-1}, and are complex motions involving two or more of the constituent molecules. Unlike the case of the dimers, it is only within these modes that the trimers exhibit large differences in frequencies between the different conformers, most notably in the H wagging modes.

Detecting these dimer and trimer species is a difficult experimental task because of their low abundance and the potential for their rovibrational spectra to overlap each other's spectra, as well as the spectra of the monomers and pure water clusters. Fortunately, there are a few modes that can be used to distinguish dimer A from trimers A and B. Table 5 summarizes

TABLE 3: Scaled Vibrational Frequencies (cm^{-1}) of HO_2, H_2O, and $HO_2\cdots H_2O$ Complex at HF/6-31G* (Scaled by 0.8929, with Experimental Values in Parentheses and IR and Raman Intensities in Brackets)[a]

mode[b]	H_2O	HO_2	dimer A	dimer B	dimer C
$\upsilon_3(H_2O)$	3740 (3756)[c] (3734)[d]		3730 [m,m] (3691)[f]	3734 [m,m]	3642 [w,s]
$\upsilon_1(H_2O)$	3634 (3657)[c] (3638)[d]		3629 [w,s] (3501.5)[f]	3633 [w,m]	3738 [w,m]
$\upsilon_1(HO_2)$		3585 (3413.0)[e]	3483 [vs,s] (3236.2)[f]	3582 [w,s]	3585 [w,s]
$\upsilon_2(H_2O)$	1631 (1595)[c]		1627 [s,w]	1643 [m,w]	1646 [m,w]
$\upsilon_2(HO_2)$		1449 (1388.9)[e]	1538 [m,w] (1479.3)[f]	1454 [w,w]	1447 [w,w]
$\upsilon_3(HO_2)$		1117 (1100.8)[e]	1126 [w,w] (1120.4)[f]	1125 [w,w]	1117 [w,w]
H wag			562 [s,0]	334 [vs,w]	312 [vs,w]
H···O stretch			332 [s,w]	91.37 [w,0]	136 [0,0]
rock (H_2O)			229 [s,w]	180 [m,w]	136 [m,w]
deformation			207 [w,0]	120 [0,0]	107 [w,w]
deformation			94 [m,w]	37 [0,0]	98 [0,0]
rock (HO_2 and H_2O)			59 [w,0]	34 [m,0]	38 [w,0]

[a] Intensities are listed as very strong (vs), strong (s), medium (m), weak (w), or negligible (0). [b] Mode motion: $\upsilon_1(H_2O)$ for OH symmetric stretching in H_2O; $\upsilon_2(H_2O)$ for HOH bending in H_2O; $\upsilon_3(H_2O)$ for OH asymmetric stretching in H_2O; $\upsilon_1(HO_2)$ for H−O stretching in HO_2; $\upsilon_2(HO_2)$ for HOO bending in HO_2; $\upsilon_3(HO_2)$ for O−O stretching in HO_2. [c] Reference 46. [d] Reference 47. [e] Reference 48. [f] Reference 49.

TABLE 4: Scaled Vibrational Frequencies (cm^{-1}) of $HO_2\cdots(H_2O)_2$ Complexes at HF/6-31G* (Scaled by 0.8929, with IR and Raman Intensities in Brackets[a])

mode[a]	trimer A	trimer B	trimer C	trimer D
$\upsilon_3(H_2O$ d)	3717 [m,m]	3721 [m,m]	3733 [w,s]	3722 [m,m]
$\upsilon_1(H_2O$ a)	3532 [s,s]	3544 [s,s]	3732 [m,s]	3561 [s,s]
$\upsilon_1(H_2O$ d)	3610 [m,s]	3613 [m,s]	3642 [w,s]	3617 [w,s]
$\upsilon_3(H_2O$ a)	3702 [m,m]	3706 [m,m]	3632 [w,s]	3709 [m,m]
$\upsilon_1(HO_2)$	3373 [vs,vs]	3379 [vs,vs]	3455 [vs,vs]	3430 [vs,s]
$\upsilon_2(H_2O$ d)	1637 [m,w]	1635 [m,w]	1652 [s,m]	1636 [w,w]
$\upsilon_2(H_2O$ a)	1649 [m,w]	1656 [m,w]	1626 [m,w]	1649 [m,w]
$\upsilon_2(HO_2)$	1568 [m,w]	1570 [m,w]	1536 [w,vw]	1512 [m,w]
$\upsilon_3(HO_2)$	1139 [w,w]	1139 [w,w]	1128 [w,vw]	1127 [w,w]
H wag	721 [m,w]	646 [m,w]	621 [m,0]	711 [w,0]
H wag	665 [m,0]	659 [m,w]	334 [m,vw]	570 [vs,w]
H wag (H_2O a, d)	455 [m,w]	403 [m,0]	297 [w,vw]	390 [s,w]
O···O stretch (H_2O a, HO_2)	263 [m,w]	240 [w,0]	234 [m,vw]	212 [w,0]
H wag	355 [m,w]	262 [m,w]	215 [m,0]	346 [m,0]
H twist (H_2O d)	246 [w,w]	384 [m,w]	140 [w,vw]	242 [w,0]
rock (H_2O a & d)	229 [w,0]	197 [w,w]	130 [w,0]	193 [m,w]
O···O stretch (H_2O a & d)	199 [w,0]	191 [w,w]	112 [w,0]	172 [w,0]
hydrogen motion	164 [m,w]	133 [m,0]	65 [w,0]	103 [m,w]
hydrogen motion	65 [w,0]	54 [w,w]	47 [w,vw]	124 [w,0]
heavy atom motion	151 [w,w]	156 [w,w]	25 [w,0]	36 [w,0]
heavy atom motion	104 [w,0]	100 [w,0]	22 [w,0]	87 [w,w]

[a] See Table 3 for description of intensities and mode motion.

TABLE 5: HF/6-31G* Scaled Infrared and Raman Absorptions for the Low Energy Clusters $HO_2\cdots(H_2O)_n$ for $n = 1$ and 2, along with the Respective Monomers[a]

modes	HO_2	H_2O	$H_2O\cdots H_2O$[b]	$H_2O\cdots(H_2O)_2$[b]	dimer A	trimer A	trimer B
υ_1 (H_2O)		3634 [m, na]	3630 [m,w]	3562 [s,m]; 3558 [s,m]; 3519 [w,s]	3629 [w,s] (−5//na)	3532 [s,s] (−102//−97)	3544 [s,s] (−90//+12)
υ_1 (HO_2)	3585 [m, na]				3483 [vs,s] (−102//na)	3373 [vs,vs] (−212//−110)	3379 [vs,vs] (−206//+6)
H motion						721 [m,w]	646 [m,w] (na//−75)
						665 [m,0]	659 [m,w] (na//−19)
			553 [m,w]	598 [s,w]	562 [s,0]	455 [m,w] (na//−107)	403 [m,0] (na//−52)
			342 [m,w]	407 [m,w]; 314 [w,w]; 296 [w,w]; 221 [w,w]		355 [m,w]	262 [m,w] (na//−93)

[a] In parentheses is given the spectral shift from the monomer//shift from the preceding cluster, where a positive value indicates a blue shift and a negative value indicates a red shift. [b] Reference 45.

this information, showing that dimer A's hydroperoxy radical's O−H stretch will be red shifted by approximately 100 cm^{-1} from that of free hydroperoxy and will have a very intense infrared peak. Trimers A and B will have this mode, further red shifted by another 100 cm^{-1}, again with very strong infrared absorption bands. Similarly, water's symmetric O−H stretch in dimer A will be unchanged from that of free water, but if trimers A or B form, a strong infrared peak will appear red shifted by approximately 100 cm^{-1}. In terms of the symmetric O−H stretching mode of the water molecule ($\upsilon_1(H_2O)$) and hydrogen wagging mode of the complex, the unique infrared peaks for dimer A and trimers A and B will be masked by vibrational modes that arise from the formation of the $H_2O\cdots H_2O$ and $H_2O\cdots(H_2O)_2$ complexes,[45] as seen in Table 5.

Conclusion

Our thorough search of conformational space discovered two unique $HO_2\cdots H_2O$ complexes in which hydroperoxy acts as a hydrogen bond acceptor, in addition to the previously described $HO_2\cdots H_2O$ complex in which hydroperoxy acts as a hydrogen bond donor. The donor complex is the most abundant of the three, resulting from a more favorable free energy of formation

for the two enantiomers of this complex. Our computed binding energies and free energies were very similar to those obtained from previous computations. Our free energy of formation is well within the error bars of the experimental value. We identified two cyclic trimers that are very similar in energy. The most favorable hydrogen-bonding motif has both hydroperoxy radical's oxygens participating in hydrogen bonding, one as a hydrogen bond donor and the other as an acceptor, with the waters forming a hydrogen bond between them. This same motif has been observed in cyclic water clusters,[27,35] and has been attributed to the enhanced cooperativity of these cyclic structures.[37] We predict that when the HO_2 concentration is on the order of 10^8 molecules·cm^{-3} in the lower troposphere at 298 K, that the number of $HO_2\cdots H_2O$ complexes is on the order of 10^7 molecules·cm^{-3} and the number of $HO_2\cdots(H_2O)_2$ complexes is on the order of 10^6 molecules·cm^{-3}.

Acknowledgment. Acknowledgment is made to the donors of the Petroleum Research Fund, administered by the American Chemical Society, to the Research Corporation, to NIH, to DOD, and to Hamilton College for support of this work. This project was supported in part by NSF Grant CHE-0457275, and by NSF Grants CHE-0116435 and CHE-0521063 as part of the MERCURY high-performance computer consortium (http://mercury.chem.hamilton.edu). T.S.D. would like to acknowledge support from the National Science Foundation under Grant ATM-0352926.

Supporting Information Available: Tables containing energy values and coordinates from the final orientation for dimer A, dimer B, dimer C, trimer A, trimer B, trimer C, trimer D, water, and HO_2 radical. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Wayne, R. P. *Chemistry of Atmospheres*, 2nd ed.; Oxford University Press: Oxford, U.K., 1991.
(2) Sander, S. P.; Peterson, M.; Watson, R. T.; Patrick, R. *J. Phys. Chem.* **1982**, *86*, 1236.
(3) Hamilton, J. E., Jr.; Lii, R.-R. *Int. J. Chem. Kinet.* **1977**, *9*, 875.
(4) DeMore, W. B. *J. Phys. Chem.* **1979**, *83*, 1113.
(5) Cox, R. A.; Burrows, J. P. *J. Phys. Chem.* **1979**, *83*, 2560.
(6) Lii, R.-R.; Sauer, J.; Myran, C.; Gordon, S. *J. Phys. Chem.* **1981**, *85*, 2833.
(7) Kircher, C. C.; Sander, S. P. *J. Phys. Chem.* **1984**, *88*, 2082.
(8) Hamilton, J. E., Jr.; Naleway, C. A. *J. Phys. Chem.* **1976**, *80*, 2037.
(9) Pan, X.; Bass, A. D.; Jay-Gerin, J.-P.; Sanche, L. *Icarus* **2004**, *172*, 521.
(10) Baragiola, R. A. *Nucl. Instrum. Methods Phys. Res. Sect. B—Beam Interact. Mater. At.* **2005**, *232*, 98.
(11) De Grey, A. *DNA Cell Biol.* **2002**, *21*, 251.
(12) Kanno, N.; Tonokura, K.; Tezaki, A.; Koshi, M. *J. Phys. Chem. A* **2005**, *109*, 3153.
(13) Aloisio, S.; Francisco, J. S.; Friedl, R. R. *J. Phys. Chem. A* **2000**, *104*, 6597.
(14) Aloisio, S.; Francisco, J. S. *J. Phys. Chem. A* **1998**, *102*, 1899.
(15) Zhu, R. S.; Lin, M. C. *Chem. Phys. Lett.* **2002**, *354*, 217.
(16) Belair, S. D.; Kais, S.; Francisco, J. S. *Mol. Phys.* **2002**, *100*, 247.
(17) Lendvay, G. *Z. Phys. Chem.-Int. J. Res. Phys. Chem. Chem. Phys.* **2001**, *215*, 377.
(18) Zhu, R. S.; Lin, M. C. *Physchemcomm* 2001.
(19) Zhu, R. S.; Lin, M. C. *Physchemcomm* **2003**, *6*, 51.
(20) Shi, Q. C.; Belair, S. D.; Francisco, J. S.; Kais, S. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 9686.
(21) Sennikov, P. G.; Ignatov, S. K.; Schrems, O. *Chemphyschem* **2005**, *6*, 392.
(22) Hansen, J. C.; Francisco, J. S. *Chemphyschem* **2002**, *3*, 833.
(23) SPARTAN. Spartan; 5.1.3 ed.; Wavefunction, Inc.: Irvine, CA 92612, 1998.
(24) Day, M. B.; Kirschner, K. N.; Shields, G. C. *J. Phys. Chem. A* **2005**, *109*, 6773.
(25) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*, revision C.02. Gaussian, Inc.: Wallingford, CT, 2004.
(26) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Rassolov, V.; Pople, J. A. *J. Chem. Phys.* **1998**, *109*, 7764.
(27) Dunn, M. E.; Pokon, E. K.; Shields, G. C. *J. Am. Chem. Soc.* **2004**, *126*, 2647.
(28) Pickard, F. C.; IV; Dunn, M. E.; Shields, G. C. *J. Phys. Chem. A* **2005**, *109*, 4905.
(29) Pickard, F. C.; IV; Pokon, E. K.; Liptak, M. D.; Shields, G. C. *J. Chem. Phys.* **2005**, *024302*.
(30) Tsai, C. J.; Jordan, K. D. *J. Phys. Chem.* **1993**, *97*, 5208.
(31) Xantheas, S. S. *J. Chem. Phys.* **1994**, *100*, 7523.
(32) Xantheas, S. S. *J. Chem. Phys.* **1995**, *102*, 4505.
(33) Gruenloh, C. J.; Carney, J. R.; Arrington, C. A.; Zwier, T. S.; Fredericks, S. Y.; Jordan, K. D. *Science* **1997**, *276*, 1678.
(34) Ugalde, J. M.; Alkorta, I.; Elguero, J. *Angew. Chem., Int. Ed.* **2000**, *39*, 717.
(35) Dunn, M. E.; Pokon, E. K.; Shields, G. C. *Int. J. Quantum Chem.* **2004**, *100*, 1065.
(36) Day, M. B.; Kirschner, K. N.; Shields, G. C. *Int. J. Quantum Chem.* **2005**, *102*, 565.
(37) Xantheas, S. S. *NATO ASI Series, Ser. C: Math. Phys. Sci.* **2000**, *561*, 119.
(38) Keutsch, F.; Saykally, R. J. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 10533.
(39) Xantheas, S. S.; Burnham, C. J.; Harrison, R. J. *J. Chem. Phys.* **2002**, *116*, 1493.
(40) Ohno, K.; Okimura, M.; Akai, N.; Katsumoto, Y. *PCCP Phys. Chem. Chem. Phys.* **2005**, *7*, 3005.
(41) Heard, D. E.; Pilling, M. J. *Chem. Rev.* **2003**, *103*, 5163.
(42) Milligan, D. E.; Jacox, M. E. *J. Chem. Phys.* **1963**, *38*, 2627.
(43) Yamada, C.; Endo, Y.; Hirota, E. *J. Chem. Phys.* **1983**, *78*, 4379.
(44) Thompson, W. E.; Jacox, M. E. *J. Chem. Phys.* **1989**, *91*, 3826.
(45) Dunn, M. E.; Evans, T. M.; Kirschner, K. N.; Shields, G. C. *J. Phys. Chem. A* **2005**, *110*, 303.
(46) Fraley, P. E.; Narahari-Rao, K. *J. Mol. Spectrosc.* **1969**, *29*, 348.
(47) Engdahl, A.; Nelander, B. *J. Mol. Struct.* **1989**, *193*, 101.
(48) Jacox, M. E. *J. Phys. Chem. Ref. Data* **1994**, 1.
(49) Nelander, B. *J. Phys. Chem. A* **1997**, *101*, 9092.